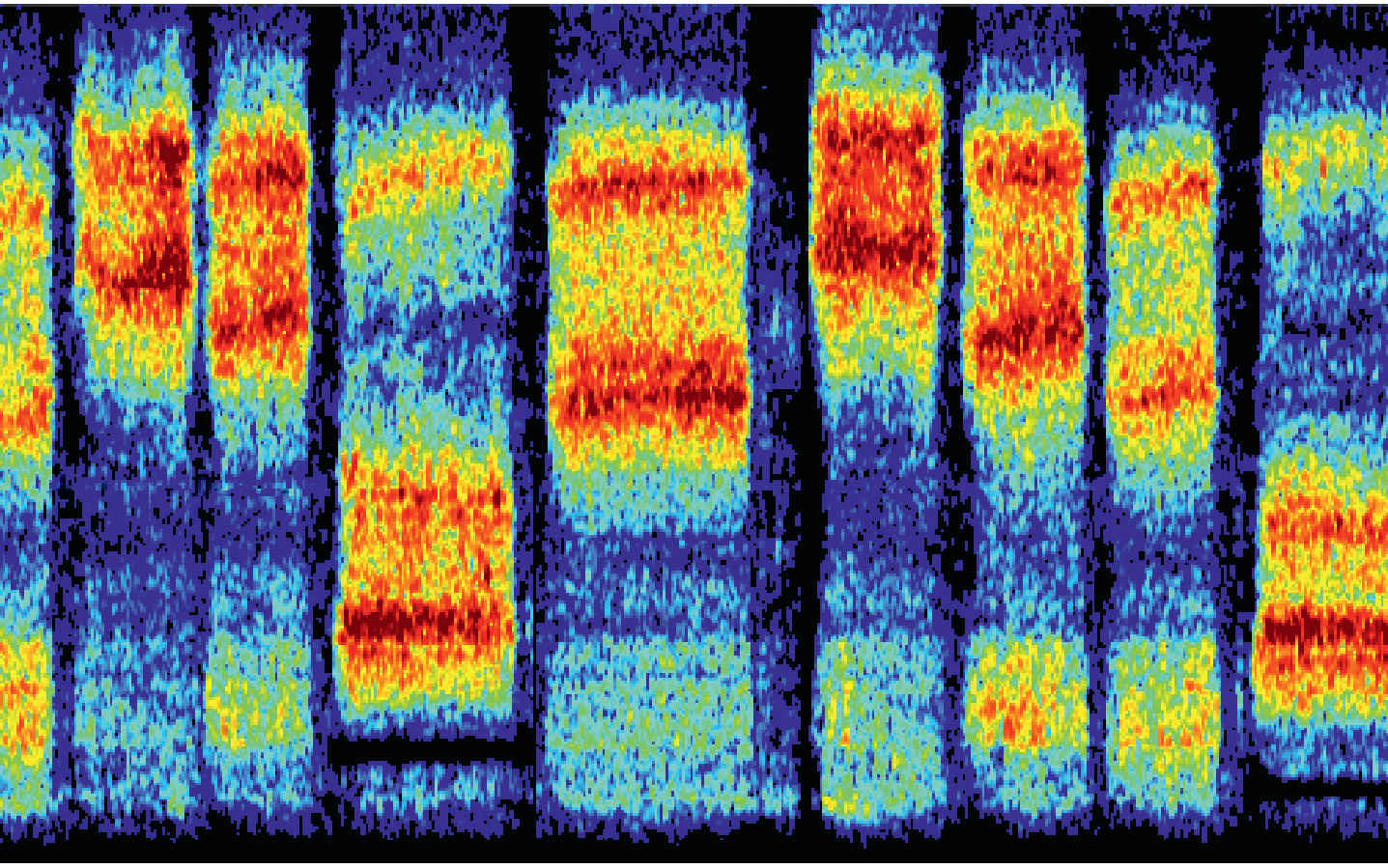
Michèle **Castellengo**

ÉCOUTE MUSICALE ET ACOUSTIQUE

Avec 420 sons et leurs sonagrammes décryptés



Sur le DVD-Rom d'accompagnement

- Des « livrets-sons » au format ePub pour l'écoute sur Smartphone
- Le livre complet en PDF avec les 420 sons intégrés

EYROLLES

ÉCOUTE MUSICALE ET ACOUSTIQUE

Point d'entrée de cet ouvrage, les 420 exemples sonores fournis sur le DVD-Rom placent le lecteur dans une situation d'écoute attentive, à partir de laquelle il va découvrir au fil des chapitres comment analyser la structure acoustique des sons et évaluer leurs qualités musicales.

L'analyse s'appuie sur l'usage du sonagramme, représentation visuelle qui révèle la structure spectrale et temporelle des sons, et se lit à la manière d'une partition musicale, sans nécessiter de connaissances avancées en acoustique.

Mais l'analyse spectrotemporelle ne suffit pas pour rendre compte des qualités d'un son dont l'écoute diffère d'une personne à l'autre. L'auteur fournit ainsi une synthèse des données récentes sur le système auditif, la reconnaissance des formes et la catégorisation cognitive qui permettent ensuite d'aborder l'étude des qualités musicales des sons – intensité, hauteur, timbre –, les problèmes posés par l'accordage des sons instrumentaux et la grande diversité des écoutes de la voix humaine.

Après des études de musique et de musicologie, Michèle Castellengo rejoint le Laboratoire d'acoustique musicale (LAM) d'Émile Leipp où elle soutient une thèse sous sa direction. En 1982, elle entre au CNRS et prend la direction du LAM. Ses recherches portent sur l'acoustique des flûtes, de l'orgue, de la voix chantée et, plus généralement, sur la perception des sons musicaux. Elle crée en 1989 la classe d'acoustique musicale du Conservatoire nationale supérieur de musique et de danse de Paris et dirige le master Atiam (Paris 6/Ircam/SupTélécom) de 1999 à 2003. Elle est aujourd'hui directrice de recherche émérite au CNRS.

Contenu du DVD-Rom d'accompagnement

- Des « livrets-sons » au format ePub 3 (un par chapitre) pour écouter les sons sur son Smartphone ou sa tablette.
- Le livre complet au format PDF avec les 420 sons aux formats MP3 et WAV.

CONFIGURATION NÉCESSAIRE. Pour les livrets-sons au format ePub : Appli iBooks pour iPod (version 4 ou ultérieure), iPhone (version 3g ou ultérieure) ou iPad (version 2 ou ultérieure), avec iOS 6 ou version ultérieure – Appli Gitden Reader (gratuite) ou Moon+ Reader (gratuite) pour Smartphones ou tablettes Android, avec Android 4 ou version ultérieure. Pour le livre au format PDF : sur Mac/PC : Adobe Acrobat Reader 6 ou version ultérieure, Mac OS X 10.0 ou version ultérieure, Windows XP ou version ultérieure – Sur iPad (version 2 ou ultérieure) : applis payantes PDF Expert (9,99 €) ou ezPDF Reader (3,99 €), iOS 6 ou version ultérieure – Sur tablettes Android : appli payante ezPDF Reader (3,22 €), Android 4 ou version ultérieure.

www.editions-eyrolles.com
Groupe Eyrolles | Diffusion Geodif

Studio Eyrolles © Éditions Eyrolles

Code éditeur : G13872
ISBN : 978-2-212-13872-6

« Avec l'exigence d'une musicienne et la rigueur d'une scientifique ouverte à toutes les musiques et voix du monde, l'auteur réunit dans cet ouvrage le fruit d'une vie consacrée à la recherche et à l'enseignement.

Ses connaissances, autant que la manière de les partager, se sont enrichies et affinées au contact des musiciens, des acousticiens, des linguistes, des luthiers, des compositeurs et des ethnomusicologues, qui viennent chercher dans l'acoustique musicale des réponses à leurs questions.

Par la richesse et la diversité des exemples sonores qu'il contient, ce livre invite à de fascinantes expériences d'écoute, où le plaisir de la découverte accompagne la satisfaction de comprendre : il ravira autant les musiciens et mélomanes curieux que les passionnés du son.

Hugues Genevois, directeur du LAM

ÉCOUTE MUSICALE ET ACOUSTIQUE

Michèle **Castellengo**

ÉCOUTE MUSICALE ET ACOUSTIQUE

Avec 420 sons et leurs sonagrammes décryptés

Préfaces de Jean-Sylvain Liénard et Georges Bloch

EYROLLES

Crédits iconographiques

Sauf mention particulière, tous les sonagrammes ont été réalisés par l'auteur à l'aide du logiciel AudioSculpt de l'Ircam.

Les sources des illustrations qui n'ont pas été réalisées par l'auteur elle-même sont mentionnées dans leur légende, à l'exception des schémas des figures 6, 7 et 1.1, réalisés par Antoine Moreau-Dusault.

L'éditeur a fait tout son possible pour identifier les ayants droit des visuels présentés. Si toutefois l'un d'eux avait été oublié, il est invité à se mettre en contact avec les Éditions Eyrolles.

Source et copyright des sons

Le concept de ce livre repose sur la fourniture de nombreux exemples choisis parmi des œuvres existantes ou réalisés spécialement pour permettre une écoute riche en contexte musical. La source des séquences sonores est indiquée entre crochets dans la section *Les sons du chapitre x* placée à la fin de chaque chapitre. Par exemple :

> **Son 6.27** – Guimbarde et chant harmonique de style sygyt. Successivement : jeu de la guimbarde (fondamental à 88,5 Hz) ; 13'' guimbarde et voix (à l'octave supérieure de la guimbarde) ; 30'' deuxième séquence voix et guimbarde ; 45'' guimbarde seule. *Tuva, Voices from the Center of Asia* ; n° 16 ; Smithsonian Folkways Records, 1990. [SF 40017]

Toute reproduction ou représentation de ces sons est interdite sans l'accord de leurs ayants droit.

De nombreux sons proviennent de la base de données RWC Music Database. Les demandes d'autorisation de reproduction de ces sons peuvent être faites directement sur le site https://staff.aist.go.jp/m.goto/RWC-MDB/.

Les sons créés par l'auteur, référencés [M. C.], sont utilisables librement pour l'enseignement et la recherche (cours, conférences) à condition d'en citer la source sous la forme suivante :

> Extrait du DVD-Rom d'accompagnement de l'ouvrage *Écoute musicale et acoustique* de Michèle Castellengo, Éditions Eyrolles, 2015.

Pour tout autre usage, il est nécessaire d'obtenir l'autorisation de l'auteur (michele.castellengo@upmc.fr).

Aux termes du Code de la propriété intellectuelle, toute reproduction ou représentation intégrale ou partielle de la présente publication, faite par quelque procédé que ce soit (reprographie, microfilmage, scannérisation, numérisation…) sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite et constitue une contrefaçon sanctionnée par les articles L.335-2 et suivants du Code de la propriété intellectuelle. L'autorisation d'effectuer des reproductions par reprographie doit être obtenue auprès du Centre français d'explitation du droit de copie (CFC) – 20, rue des Grands-Augustins – 75006 Paris.

© Groupe Eyrolles, 2015
ISBN : 978-2-212-13872-6

ÉDITIONS EYROLLES
61, bd Saint-Germain
75240 Paris Cedex 05
www.editions-eyrolles.com

Préfaces

Jean-Sylvain Liénard

Directeur de recherche émérite au CNRS

Ce livre n'est pas un livre comme les autres. C'est un support écrit et sonore, destiné à guider le lecteur auditeur dans la découverte d'un paysage infini où se mêlent musique et acoustique. Michèle Castellengo présente ici la quintessence d'une cinquantaine d'années de recherche passionnée, originale et exigeante, sur la structure, la production et la perception des sons. Son oreille exceptionnelle, sa curiosité pour le monde auditif et la qualité de son travail expérimental l'ont depuis longtemps amenée à analyser tous les sons qu'elle rencontre. Elle est aidée en cela par une méthode de travail et par un outil d'analyse, le *sonagraphe*, qui permet de représenter le son comme le fait une partition musicale, avec le temps en abscisse et la fréquence – hauteur – en ordonnée. Mieux qu'une partition, le sonagramme indique également le timbre du son. En figeant le temps, il permet d'en examiner à loisir les moindres détails et, par des réglages adéquats, d'en révéler les structures d'une manière proche de la perception naturelle.

La méthode de travail est celle qu'a inaugurée Émile Leipp, fondateur et responsable du LAM, Laboratoire d'acoustique musicale, dans les années 1960-1980. Elle repose sur l'idée que les instruments de musique, élaborés au fil du temps par les luthiers, représentent un compromis optimum entre les capacités sensorimotrices des exécutants, les capacités perceptives des auditeurs, les techniques de lutherie disponibles et les canons esthétiques d'une société et d'une époque. Pour comprendre comment fonctionne un instrument il faut donc, en tout premier lieu, prendre en compte ce que les praticiens, facteurs et instrumentistes en disent, même lorsque les règles de leur art se sont fondues dans une tradition qui n'est arbitraire qu'en apparence. Ce n'est qu'après que l'on peut, par l'analyse, en découvrir les raisons, souvent extraordinairement fines et pertinentes, et éventuellement contribuer à résoudre les problèmes posés par l'évolution continue des techniques de lutherie et des styles musicaux. Une telle doctrine implique une ouverture sur diverses disciplines, en particulier la physique, la psychologie, la musique, pour n'en citer que quelques-unes. C'est pourquoi le LAM a d'emblée rassemblé un auditoire et des collaborateurs venant de tous horizons.

Michèle Castellengo a été, dès l'origine, la cheville ouvrière du laboratoire. Elle est rapidement devenue experte dans l'utilisation et les réglages du *sonagraphe*, à une époque où l'acoustique s'intéressait plutôt aux sons fixes et à leur représentation fréquence-amplitude. Sa spécialité a toujours été de se fier d'abord à son écoute pour repérer les phénomènes intéressants sur le plan musical ou perceptif, puis de chercher la meilleure manière de les mettre en évidence dans les sonagrammes. Elle a ainsi analysé diverses classes de sons : la flûte, le clavecin, le piano et l'orgue ; la parole, la voix et le chant ; les sons de musiques pratiquées dans d'autres cultures que la nôtre ; les chants d'oiseaux et les scènes sonores que l'on rencontre dans la vie de tous les jours. Ces études, souvent menées à l'occasion de rencontres avec des praticiens de la musique, constituent un ensemble d'où émerge une vision cohérente du monde sonore, allant de sa production par une source humaine ou physique à sa perception par l'auditeur. C'est cette vision d'ensemble qui nous est restituée dans le livre de manière pédagogique et intuitive.

Dans les années 1960, l'enthousiasme partagé par les chercheurs du LAM pour la représentation sonagraphique les a conduits à imaginer l'opération inverse : passer du document sonagramme au son correspondant. Cet appareillage, baptisé *Icophone*, construit au Laboratoire de mécanique de l'université Paris VI, permettait de transformer instantanément en sons les sonagrammes schématiques tracés à la main sur une bande transparente. Le son résultant n'était pas très harmonieux, mais la parole ainsi reproduite était compréhensible, pour peu que les schémas respectent l'évolution temporelle des structures acoustiques du signal original. À côté de développements inspirés par la théorie de la forme, il est apparu qu'on pouvait découper le flux de parole en éléments allant d'un son au suivant de façon à respecter les transitions, essentielles pour l'intelligibilité. Michèle Castellengo a mis au point expérimentalement un dictionnaire d'environ 600 éléments couvrant la langue française, permettant de construire des phrases nouvelles par assemblage à la manière des dominos. La méthode, requérant peu de mémoire, se prêtait à une mise en œuvre informatique, réalisée ultérieurement au LIMSI, Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur, laboratoire de mécanique des fluides nouvellement installé sur le campus d'Orsay. Par la suite le LIMSI a continué dans la voie du traitement automatique de la parole. Le thème s'est élargi à l'ensemble de la communication homme-machine, qui occupe aujourd'hui plusieurs centaines de chercheurs, enseignants et doctorants. Seuls les plus anciens savent que cette activité majeure du laboratoire a démarré au LAM en 1965 à partir d'une curieuse machine à inverser les sonagrammes. Ainsi va la recherche…

Le terme de parole désigne ce qui, dans le signal oral, porte l'information linguistique. La voix en est le support physique et, bien entendu, les deux notions sont fortement dépendantes l'une de l'autre. La voix chantée joue un rôle primordial en musique, dans toutes les cultures. Dans la musique classique européenne, tous les apprentis chanteurs sont confrontés au problème du passage, discontinuité de timbre qui s'entend quand on produit un glissando allant de l'extrême grave à l'extrême aigu. Les passages sont dus à des transitions entre plusieurs régimes vibratoires des cordes vocales. Quatre régimes vibratoires distincts, appelés mécanismes, ont été mis en évidence au LAM et cette notion fondamentale commence à être reconnue dans la communauté scientifique de la voix. Michèle Castellengo a également mené ou encadré des études sur d'autres aspects de la voix chantée, toujours en relation étroite avec des chanteurs professionnels. Elle s'est aussi intéressée au chant multiphonique que l'on trouve dans certaines cultures populaires ou extra-européennes ; les exemples spectaculaires présentés au chapitre 9 témoi-

gnent à la fois du talent des artistes et de la pertinence des explications obtenues au moyen de l'analyse sonagraphique.

Un autre point à souligner est l'intérêt porté aux processus perceptifs et cognitifs mis en œuvre dans l'activité d'écoute. La psychoacoustique est une discipline à part entière, qui utilise une méthodologie rigoureuse, avec des sons calibrés et des tests validés statistiquement. Elle semble donc très loin de l'écoute individuelle de fines nuances musicales. Pourtant Michèle Castellengo cherche en permanence à interpréter ses résultats en fonction des connaissances acquises sur la perception auditive. Certaines de ses études conduisent à poser des questions de nature psychoacoustique. Comment se fait-il, par exemple, que la zone fréquentielle allant approximativement de 500 à 1 700 Hz s'avère si importante pour la perception du contenu des sons, alors que l'on sait depuis longtemps que le maximum de sensibilité de l'oreille se trouve aux alentours de 3 000 Hz ? Dans la dimension temporelle, il est fascinant de constater que, selon la cadence de répétition d'une simple impulsion sonore, la sensation produite va d'une suite de clics distincts à un son doté d'une texture rythmique sans hauteur, puis à un son complexe pourvu d'une hauteur, et enfin à un son très aigu sans sensation de hauteur. Des observations de ce genre, traitées empiriquement par les musiciens à des fins artistiques, se posent en permanence en acoustique musicale et constituent des objets d'étude encore peu explorés en psychoacoustique.

Cet ouvrage arrive à un moment où le son occupe une place inédite dans notre société : sons de la radio, sons de la rue, bruit, parole, musique, télévision, enregistrement, transmission, sonorisation, insonorisation, communication, malentendance : le son est partout. Les nouveaux outils numériques permettent de produire tous les sons imaginables, mais, pour en faire bon usage, musiciens et designers sonores devront être en mesure de choisir ceux qui correspondent à leurs besoins précis. Pour ces créateurs, pour les concepteurs de systèmes sonores et pour les « ingénieurs en sensations » dont Abraham Moles avait prévu l'arrivée avec trente ans d'avance, les connaissances exposées dans ce livre s'avéreront rapidement indispensables. Pour les autres lecteurs, amoureux du son, de la musique et de la science, elles seront une double source de joie : joie de l'esprit, par l'analyse et la compréhension, et joie sensorielle, par l'écoute et la découverte d'un monde sonore sans limite.

Georges Bloch

Compositeur et chercheur

C'était dans les années 1980 : les quelques étudiants français du département de musique de l'université de Californie à San Diego le surnommaient « le fils d'Éric ». C'était bien peu respectueux pour un professeur non seulement admiré pour sa culture et son inventivité musicale, mais qui, de plus, souffrait d'une forme rare de myopathie qui l'obligeait à se déplacer en chaise roulante. Mais Robert Erickson (1917-1997) était un monsieur spécial et son séminaire était un passage obligé à UCSD. On aurait pu l'appeler un séminaire d'orchestration, si le terme n'avait eu encore à l'époque une connotation très post-berliozienne ; ou « séminaire sur le timbre », si « timbre » n'était pas devenu un terme tellement polysémique qu'on ne savait plus de quoi il s'agissait exactement.

Chez Erickson, on savait de quoi on parlait : il avait un point de vue, qu'il ne nous obligeait absolument pas à partager, mais qui permettait de savoir sur quelles prémisses se fondait sa recherche ; et il avait une volonté encyclopédique d'explorer toutes les voies par lesquelles un compositeur peut imaginer des structures musicales à partir du son lui-même. De ses réflexions était sorti un livre, *Sound Structures in Music,* peut-être le seul « traité d'orchestration » intéressant du XXe siècle.

Hélas, dans ce livre, il y a un problème : il n'y a pas de son. Pour ceux qui ont eu la chance d'assister à son cours, aucune importance : nous passions notre temps à écouter des enregistrements, à analyser des partitions en détail, et des interprètes venaient faire des démonstrations. Quand on a juste le livre, c'est plus difficile, car les copies de partitions d'orchestre réduites à la taille d'un bouquin deviennent vite illisibles ; et puis, justement, tout le son n'est pas sur la partition. Or, le travail d'Erickson était directement issu de sa compétence comme compositeur et, aussi, de sa curiosité : dès qu'il découvrait une combinaison sonore qui l'étonnait, il voulait savoir comment cela marchait.

L'ouvrage de Michèle Castellengo a de nombreux points communs avec celui d'Erickson. Tout d'abord, il résulte d'une longue expérience d'écoute, plus précisément d'écoute *curieuse*. En revanche, il possède un grand atout que n'avait pas son

prédécesseur. Il commence par le début : les sons y sont. Rien que pour cela, ce « livre » est incroyable, inestimable : écoutez ces sons inouïs ; après, vous aurez envie d'en savoir plus. On peut même aller plus loin : le simple fait d'écouter ces sons rend savant, simplement parce que ce sont presque tous des sons fascinants, captivants, et parce qu'ils questionnent l'écoute.

Ce point de départ – qui est aussi un point d'arrivée – ne vaut pas que pour le lecteur : c'est aussi le point de départ de l'auteur. En effet, comme celui d'Erickson, ce livre s'appuie sur un point de vue, clairement exprimé dans le quatrième chapitre, celui concernant la perception. Ce point de vue, dit écologique, part du « sujet connaissant », et surtout de l'expert de la pratique du son. Michèle Castellengo a profité de la myriade de savants qui ont fréquenté le Laboratoire d'acoustique musicale (LAM) de Jussieu : des acousticiens, bien sûr, mais aussi des instrumentistes, des luthiers, des linguistes, des compositeurs. Essentiellement des gens dont l'écoute est le métier et la passion. Comme disait le « fils d'Éric » dans son séminaire : « tous les clarinettistes distinguent les changements de registre ; allez donc voir un clarinettiste et demandez-lui de vous les jouer et de vous les expliquer jusqu'au moment où, vous aussi, vous les entendrez, y compris lorsqu'on les masque ».

On part donc du son, et, plus précisément, de l'écoute experte de ce son. Il y a les « oreilles d'or » de la marine, il y a « l'écoute critique » des preneurs de son, l'écoute spécialisée de ceux qui font de la musique électroacoustique, et, bien entendu, l'écoute qu'acquièrent les instrumentistes à la suite de la pratique quotidienne de leur instrument. C'est cela qui intéresse Michèle Castellengo : comment on écoute, qu'est-ce qu'on écoute, et comment vous, moi, Pierre-Yves Artaud, M. Dupond, avec sa compétence – ou son ignorance – de musicien, de luthier, de soudeur, de chauffeur de taxi, de berger jouant du cor des Alpes, d'amateur de musique baroque, de spécialiste de techno, projette son savoir sonore et ses représentations sur ces sons. Répétons-le : c'est ce point de vue sur l'écoute qui fait toute la valeur de cet ouvrage, même si on ne le partage pas. En effet, pour citer l'auteur elle-même : « Il faut mettre en garde le lecteur sur le fait que des termes comme *perception*, *information*, *forme*, revêtent des sens et des contenus souvent fort différents d'un ouvrage à l'autre. » Il y a un point de vue revendiqué ? Au moins, on sait de quoi on parle.

Le LAM de Michèle Castellengo, c'est aussi une aventure de plus de cinquante ans, et des dizaines de sujets de recherche, dont beaucoup se retrouvent dans ce livre. Cela donne un ouvrage dense : en vingt-cinq pages, on traite de la voix humaine, depuis les traités de chant du XIXe siècle et les articles fondateurs de Sundberg sur les formants jusqu'aux derniers travaux de Nathalie Heinrich sur l'effet des rapports entre les vibrations de la glotte et du larynx. Vingt-cinq pages, c'est peu pour beaucoup de chose, surtout avec tous ces sons merveilleux : la *quintina* sarde, divers types de chant harmonique, les résonances buccales de guimbarde. De toute façon, les sons y sont.

Un ouvrage si compact pourrait se résumer à une sorte de dictionnaire raisonné de l'écoute musicale de l'acoustique – ce qui ne serait déjà pas si mal, d'ailleurs. Mais c'est bien plus, car l'expérience pédagogique de Michèle Castellengo, notamment à la classe d'acoustique musicale du Conservatoire de Paris, fait que la présentation des questions est toujours passionnante. Les questions du timbre, de la perception des hauteurs et des intensités restent ce qu'elles sont : des questions, qu'il convient de préciser en fonction des expériences sonores. Pendant de nombreuses années,

l'auteur a réussi à faire dialoguer divers spécialistes de l'écoute, en essayant de leur faire cracher ce qu'ils entendaient lorsqu'ils maniaient les sons. C'est une pédagogie qui n'était pas désintéressée : toute sa recherche partait de là, de ces écoutes expertes !

Dans ce livre, les processus d'écoute des experts, qui produisent, travaillent ou utilisent le son, sont décortiqués ; c'est au lecteur de s'y mesurer. Parmi les exemples sonores, il y en a aussi quelques-uns volontairement banals, comme le premier, celui qui ouvre l'introduction ; on finit cependant par leur accorder autant d'attention qu'aux extraits sonores plus inouïs de cet ouvrage – et ils sont nombreux ! On veut comprendre : les sons y sont.

Avant-propos

Voici un livre à lire et à écouter qui offre plus de 400 exemples sonores patiemment récoltés au long de décennies de recherches en acoustique sur les sujets les plus divers : des flûtes à l'orgue, des cloches aux oiseaux, de la parole synthétique au chant lyrique ou traditionnel. En associant l'audition de sons qui réjouissent les oreilles, leur visualisation et la lecture de données explicatives, notre ambition est d'introduire dans l'univers de l'acoustique les musiciens et les auditeurs curieux, en quête de connaissances sur le monde sonore et les perceptions qu'ils en ont.

Les sons n'ont qu'une existence éphémère : à peine entendus, ils s'évanouissent, ne laissant en mémoire qu'une trace réduite sur laquelle il est difficile d'échanger des impressions comme nous pouvons le faire à la vue d'un paysage ou d'une scène photographiée. De surcroît, si plusieurs auditeurs entendent la même séquence sonore et qu'on les invite à en décrire les caractéristiques et les qualités, on constate une grande diversité de réponses : tous entendent le même son, mais chacun le perçoit et l'écoute différemment. Il faut donc capter les sons, comme l'a merveilleusement anticipé Rabelais au XVIe siècle[1] et en donner une représentation objective. L'analyse sonagraphique aujourd'hui à la portée de tous grâce à l'informatique est abondamment utilisée dans le livre, car elle possède un double avantage. C'est une visualisation du son assimilable rapidement, qui s'apparente à l'écriture musicale et qui offre la possibilité d'aborder, quand elle est couplée à l'écoute, l'étude rigoureuse des paramètres acoustiques des sons.

La démarche de l'ouvrage

La plupart des ouvrages français sont assez spécialisés et peu accessibles au commun des mortels. Les livres d'acoustique à l'exception de celui d'Émile Leipp[2] sont réservés à des lecteurs pourvus d'un bon niveau scientifique et n'abordent que rarement l'analyse des sons et leur perception ; les livres sur la musique s'intéressent peu à la matière sonore elle-même ; quant aux ouvrages traitant de la perception et de la cognition, ils donnent une large place au sens visuel, ne laissant que la portion congrue à une perception auditive de sons de laboratoire.

Nous avons donc souhaité réunir dans un seul livre les données issues des différentes disciplines qui permettent de comprendre les phénomènes sonores et leur écoute, tout en restant accessible au plus grand nombre. Les deux premiers chapitres présentent les notions élémentaires à connaître sur la production des sons, leur

1. *Le Quart Livre*, chapitre LV, « Comment en haulte mer Pantagruel ouyt diverses parolles dégelées ».
2. *Acoustique et musique*, 1971, Masson : Paris ; réédité en 2010 aux Presses des Mines.

analyse, en prenant délibérément comme exemples sonores les sons riches et intéressants de la musique. Le troisième chapitre offre une présentation schématique du système auditif et de ses capacités d'analyse du signal sonore ; le quatrième aborde à l'aide de la Gestalttheorie et de la catégorisation perceptive la perception sonore d'un auditeur confronté à différentes situations d'écoute selon qu'il entend un signal imprévisible, les sons habituels de son environnement, ceux d'une conversation dans laquelle il est engagé ou ceux de la musique. Ainsi munis de connaissances sur la structure acoustique des sons et sur nos stratégies d'écoute, nous pouvons aborder l'étude des qualités essentielles des sons musicaux : l'intensité et les caractéristiques qui font qu'un son émerge d'un ensemble (chapitre 5) ; la hauteur des diverses sortes de sons instrumentaux (chapitre 6) et la vaste question du timbre, particulièrement développée dans le chapitre 7 selon une approche nouvelle. Le chapitre 8 traite de questions spécifiques aux musiques mélodiques et harmoniques dont la dimension privilégiée est la hauteur (intervalles, systèmes d'accordage). Enfin il nous tenait à cœur d'offrir dans le dernier chapitre quelques exemples d'application de toutes ces notions à la voix humaine, l'instrument que chacun de nous possède.

On trouvera dans les annexes les conventions de notation (musicale, acoustique), quelques documentations complémentaires, une bibliographie conséquente ainsi qu'un glossaire très détaillé des termes en usage dans ce livre.

Le contenu de l'ouvrage et la pédagogie qui est à l'œuvre bénéficient de l'expérience acquise au cours d'années d'enseignements donnés à des étudiants de cursus variés : les élèves instrumentistes, compositeurs, musicologues de la classe d'acoustique musicale du Conservatoire national supérieur de musique de Paris, auxquels se joignaient aussi des ethnomusicologues et de futurs preneurs de son ; les étudiants en cinéma de la Fémis, ceux du master scientifique ATIAM (Acoustique, traitement du signal, informatique, appliqués à la musique) de l'université Paris VI et les ingénieurs de la semaine Athens « Musique, science, histoire » de l'École des mines Paris-Tech.

Un livre à écouter

L'écoute devrait, le plus souvent, précéder la découverte des analyses visuelles et de leurs commentaires. Pour aplanir l'obstacle qui persiste entre les supports de l'écrit et ceux du sonore, deux solutions sont proposées dans le DVD-Rom d'accompagnement.

La première consiste à utiliser une version numérique du livre apparaissant à l'écran d'un ordinateur (ou d'une tablette) dans la même mise en pages, mais avec les sons intégrés, donc immédiatement accessibles sur un simple clic. Cette édition PDF multimédia de l'ouvrage est proposée en deux versions : l'une avec sons au format MP3, l'autre avec sons au format WAV.

La deuxième solution s'adresse à ceux qui préfèrent la lecture sur papier. Nous leur offrons en complément du livre une série de « livrets sons », un par chapitre, contenant la collection des exemples sonores à importer dans un lecteur mobile (Smartphone ou tablette). Ces fichiers au format ePub3, réalisés par Guillaume Pellerin de la société Parisson, sont compatibles avec les plateformes Apple iOS et Android. Ils comprennent une page par son, avec la légende du son, une barre de lecture audio et une vignette rappelant la figure associée. Vous pouvez également accéder direc-

tement à ces livrets sons en scannant le QR code ci-contre à l'aide de votre Smartphone ou de votre tablette, ou via l'URL : www.editions-eyrolles.com/go/castellengo.

Toutes les indications sur l'usage de ces documents sonores sont fournies dans l'annexe « Contenu du DVD-Rom d'accompagnement », en fin d'ouvrage.

Quelle modalité d'écoute adopter ? L'écoute des sons sortant directement d'un Smartphone, d'une tablette ou d'un ordinateur est une écoute de contrôle de qualité très réduite, très limitée vers les basses fréquences. Elle est à prohiber.

L'écoute au casque, très répandue aujourd'hui, a le double avantage d'être discrète et d'isoler du bruit environnant. Elle convient lorsque le casque est de bonne qualité, mais nous recommandons de l'éviter pour une première approche, car elle nous prive d'une fonction essentielle, l'exploration du champ sonore par de petits mouvements de tête pour apprécier les différences entre les sons arrivant aux deux oreilles, ressource importante pour la discrimination qualitative des sons.

L'idéal est donc de pouvoir écouter fréquemment sur une bonne chaîne de reproduction, si possible en compagnie d'auditeurs amis impliqués dans des pratiques sonores différentes afin de découvrir, au cours d'échanges réciproques, l'étonnante richesse des écoutes individuelles.

À Émile Leipp,
créateur du LAM
et à Abraham Moles,
passeur d'idées

À Marie-José,
par qui tout est arrivé

Remerciements

Un tel ouvrage n'aurait pu voir le jour sans l'engagement de chercheurs passionnés par le projet qui ont, pendant près de dix années, tenu un séminaire de suivi de la rédaction : Denis Mercier, créateur sonore cinéma et musique, directeur de l'ouvrage collectif *Le livre des techniques du son*, initiateur et aiguilleur du projet ; Pascal Gaillard, musicologue, chercheur en perception auditive et maître de conférences à l'université de Toulouse ; Charles Besnainou, luthier et ingénieur de recherche en acoustique instrumentale au LAM ; Marie-Cécile Barras, musicologue, maître de conférences à l'université de Bordeaux ; Thierry Maniguet, musicologue, conservateur au musée de la Musique ; Hugues Genevois, chercheur en nouvelles technologies et création musicale, ingénieur de recherche au ministère de la Culture, responsable de l'équipe LAM ; Adrien Mamou-Mani, concepteur d'instruments augmentés, chercheur à l'Ircam et professeur d'acoustique musicale au Conservatoire national supérieur de Paris. Tous ont contribué – par leurs compétences complémentaires – à l'élaboration de l'ouvrage en débattant des idées proposées et en fournissant un soutien stimulant jusqu'à l'aboutissement de ce livre. Ma dette envers eux est immense. S'y ajoute celle que j'ai envers Danièle Dubois dont les idées ont irrigué et conforté mon approche intuitive de l'écoute en m'initiant à la catégorisation perceptive.

La collection des sons, qui fait la richesse de cet ouvrage, provient en grande partie des archives du LAM[1] constituées au cours des recherches développées avec Émile Leipp. Je remercie très sincèrement les nombreux instrumentistes, chanteurs, ethnomusicologues qui m'ont autorisée à reproduire leurs exemples sonores.

1. En cours de numérisation, les archives sonores du LAM sont consultables sur le site http://telemeta.lam.jussieu.fr/

D'autres proviennent de missions effectuées pour le ministère de la Culture (orgues, cloches), de séminaires et de travaux de thèse. Ma reconnaissance va également aux personnes et aux institutions qui m'ont accordé l'autorisation de reproduire les tests auditifs et les sons paradoxaux propres à questionner l'écoute musicale, ainsi qu'aux personnes par l'entremise desquelles certaines demandes ont pu aboutir : Johanni Curtet, Laurent Daudet, Jean During, Jean Kergomard, Ted Levin, Jean-Claude Risset et Joe Wolfe. Toutes les sources sont mentionnées explicitement dans le texte des légendes des sons afférentes à chaque chapitre.

Nous avons donné la priorité aux sons les plus intéressants, parfois enregistrés dans des conditions difficiles. Vincent Mons, que nous remercions vivement, a généreusement assuré le long travail de toilettage et de montage des 422 exemples sonores.

Mes remerciements vont aussi aux nombreux « relecteurs écouteurs » mis à contribution sur diverses parties relevant de leur expertise : Pierre-Yves Asselin, Daniel Fargue, Joël Frelat, Claudia Fritz, Suzanne Fürniss, Nathalie Henrich-Bernardoni, Sylvain Lamesch, Benoît Navarret, Marc Pinardel, Jean-Dominique Polack, Laurent Quartier, Corsin Vogel.

La rédaction finale a fait l'objet d'une relecture intégrale, exigeante et attentive de la part de deux personnalités représentatives des deux domaines – science et musique – ici réunis : Jean-Sylvain Liénard et Georges Bloch. Ils m'ont fait l'amitié d'écrire chacun une préface et je les en remercie très profondément.

Enfin, pour que l'ensemble de ce travail prenne la forme d'un livre, il a fallu la complicité de Daniel Fargue et Béatrice Avakian, la confiance et l'engagement d'Éric Sulpice, directeur éditorial des Éditions Eyrolles, qui a pris le risque d'un « livre-sons » particulièrement complexe à réaliser, et la formidable implication de Françoise Barat et de toute l'équipe des Éditions Eyrolles. Ce projet a bénéficié de la disponibilité qu'offre la position de chercheur émérite au CNRS. Accueillie à l'institut Jean-Le-Rond-d'Alembert (UPMC-Sorbonne Universités), j'ai pu poursuivre le développement de mon travail au sein du LAM – équipe issue du Laboratoire d'acoustique musicale créé par Émile Leipp en 1963 –, lieu où s'opère une alchimie originale entre musiciens, scientifiques, luthiers et chercheurs en sciences humaines.

Table des matières

INTRODUCTION AU MONDE DES SONS 1

Les sons de l'introduction 6

CHAPITRE 1 – **DES VIBRATIONS AUX SONS DE LA MUSIQUE** 7

1. À l'origine du son : le mouvement 7
2. La production et la propagation des sons 7
   - 2.1. En bref 7
   - 2.2. L'excitation : impulsion ou entretien 9
   - 2.3. La notion d'onde 12
   - 2.4. Les modes vibratoires (transition ondes-modes) 14
   - 2.5. Des vibrations de la structure au son rayonné 17
   - 2.6. La réception, la saisie du son, les transformations de la vibration 18
   - 2.7. Définitions utiles 19
3. La forme temporelle des vibrations 19
   - 3.1. L'inscription des vibrations 19
   - 3.2. La composition des vibrations 20
   - 3.3. Du mouvement à l'onde sonore 20
4. L'analyse auditive des composantes d'un son : la série harmonique 21
   - 4.1. L'analyse auditive par filtrage 21
   - 4.2. Le verre percuté et le verre frotté 21
   - 4.3. La série harmonique 22
   - 4.4. Les intervalles de la série harmonique 22
5. Vocabulaire : de la musique à l'acoustique et inversement 24
   - 5.1. Les divers types de sons 24
   - 5.2. Les termes à connaître : harmoniques, partiels, fondamental 24
6. Les mélodies spectrales et les mélodies de partiels successifs : exemples musicaux 26
   - 6.1. Les mélodies spectrales : sélection d'harmoniques par résonance buccale. 26
   - 6.2. Les mélodies produites avec les partiels d'un tuyau ou d'une corde 28
7. Les sons du chapitre 1 30
   - 7.1. Types d'excitation et modes vibratoires 30
   - 7.2. Analyses auditives des composantes d'un son 31
   - 7.3. Exemples de synthèse additive numérique d'un son périodique 31
   - 7.4. Partiels et harmoniques 31
   - 7.5. Mélodies d'harmoniques : exemples musicaux 32
   - 7.6. Mélodies de partiels (tube ou corde) 32
   - 7.7. Instruments à cordes fonctionnant sur la suite des partiels 32

CHAPITRE 2 – **LA REPRÉSENTATION DES SONS** 33
1. Les analyses acoustiques et l'écoute 33
1.1. L'inscription des ondes sonores 33
1.2. La représentation des sons 34
1.3. Les étapes d'une analyse 35
2. Première exploration : une « mise en bouche » 37
3. Les paramètres du signal sonore 39
3.1. La temporalité et l'analyse des vibrations sonores 39
3.2. La mesure de la fréquence : un problème difficile 40
3.3. La transformée de Fourier à court terme et la dualité temps/fréquence 43
3.4. L'amplitude globale : l'enveloppe temporelle, la dynamique 47
3.5. Les spectres : spectres à court terme et spectres moyennés 49
3.6. Le traitement du signal et la resynthèse 51
3.7. La représentation des intervalles musicaux 52
4. Quelques exemples d'analyses 56
4.1. Un cas d'école : sonagrammes, formes d'ondes et spectres de sons électroniques 56
4.2. L'anatomie d'un son de steel-drum 58
4.3. Le verre percuté et l'entretien des deux premiers modes propres 62
4.4. L'octaviation et le quintoiement 63
4.5. La séquence complète des partiels d'une flûte et d'une clarinette 64
4.6. Une mélodie jouée à la tilinca roumaine 65
4.7. Les mélodies d'harmoniques et les mélodies de partiels 66
4.8. L'analyse de la voix humaine : formants des voyelles ou intonation ? 67
4.9. L'analyse et l'interprétation perceptive : un exemple trompeur 68
4.10. La voix chantée et le piano : comparaison de sons de hauteur fixe et de sons vibrés 69
4.11. Nuit en Vendée 69
4.12. Une séquence sonore de percussions de hauteur indéterminée 70
4.13. L'analyse d'un son de hautbois avec trois logiciels différents : Wavesurfer, Praat et Audiosculpt 71
5. En guise de conclusion 73
6. Les sons du chapitre 2 76

CHAPITRE 3 – **LE SYSTÈME AUDITIF HUMAIN** 79
1. Description sommaire de l'organe auditif 79
1.1. Introduction 79
1.2. Schéma général du système auditif 79
1.3. Les trois parties de l'oreille 82
1.4. Les voies nerveuses : de la cochlée au cortex auditif 89
1.5. Le codage de l'information sensorielle 91
2. Les caractéristiques psychophysiques de l'oreille humaine 95
2.1. Les seuils absolus 95
2.2. L'oreille considérée comme un banc de filtres : la notion de bandes critiques 98
2.3. La sensibilité aux variations d'amplitude, de fréquence et de durée 101
2.4. Réponse à l'accroissement d'une stimulation : la loi de Fechner-Weber 103
2.5. Les seuils différentiels d'intensité et de fréquence 104

2.6. La sensation d'intensité des sons purs 108
2.7. La sensation de hauteur des sons purs 110
2.8. La sensation d'intensité des sons complexes 113
2.9. La sensation de hauteur des sons complexes 113
3. Observations sur quelques phénomènes temporels 118
3.1. Du successif au continu : une frontière physiologique 118
3.2. Les variations périodiques de sons de hauteur définie : modulations et battements 120
3.3. La notion de discrimination temporelle 125
3.4. Résolution temporelle ou intégration temporelle ? 126
4. La localisation des sons et la perception de l'espace 127
4.1. La localisation et la latéralisation auditive 127
4.2. La localisation dans le plan horizontal 129
4.3. La localisation dans le plan vertical : plan sagittal (ou médian) 131
4.4. La prise en compte des pavillons 132
4.5. L'effet de précédence ou loi du premier front d'onde 132
4.6. L'estimation de la distance 132
4.7. La localisation en situation réelle 132
5. Les sons du chapitre 3 135
5.1. Simulation de surdités : filtrage de la voix parlée (voix féminine) 135
5.2. Bandes critiques, asymétrie du masquage (écoute au casque conseillée) 135
5.3. Seuils différentiels 136
5.4. Finesse de discrimination auditive 136
5.5. Variation de la sensibilité auditive avec la fréquence 136
5.6. Durée du son et sensation de hauteur tonale 136
5.7. Perception mélodique dans l'aigu et plafond du codage temporel 136
5.8. Plusieurs exemples de sons dont les composantes sont équidistantes de 250 Hz 137
5.9. Discrimination des harmoniques 137
5.10. Du discontinu au continu 137
5.11. Modulations et battements 137
6. Réponses aux tests 138

CHAPITRE 4 – **UNE APPROCHE DE LA PERCEPTION SONORE : FORMES ET CATÉGORISATION** 139
1. La perception 139
1.1. Préambule 139
1.2. Qu'est-ce que percevoir ? 140
2. À l'écoute du monde sonore environnant 143
2.1. L'audition « sens d'alerte » et l'écoute choisie 143
2.2. Deux modalités perceptives distinctes : identifier les sons ou les qualifier 144
2.3. L'acte d'écoute : sensation, mémoire, anticipation 146
2.4. L'étude « écologique » de la perception sonore et le paradoxe de la complexité 149
3. La notion de forme sonore spectrotemporelle : sources et séquences 152
3.1. Qu'est-ce qu'une forme ? 152
3.2. Les formes fortes et les formes faibles 153
3.3. Peut-on parler de formes sonores ? Expérience de reconnaissance de sons très brefs 154

3.4. Typologie acoustique des formes sonores 157
3.5. Les formes sonores : primauté de la dimension temporelle 158
3.6. Les formes et le matériau sonore : deux niveaux de structuration temporelle des formes acoustiques 159
3.7. Les formes sonores et la variabilité 166
3.8. Les rapports entre la forme et le fond 169
3.9. La musique et la théorie de la forme : l'organisation perceptive 173
3.10. Les ambiguïtés d'écoute et les illusions 180
3.11. La transformation d'une forme dans une autre : le morphing 180
4. La catégorisation perceptive des sons et des séquences 182
4.1. La notion de catégorie 182
4.2. La catégorisation prototypique : similarités, niveau de base et typicalité 183
4.3. La catégorisation et la perception sonore 185
4.4. La catégorisation libre associée aux verbalisations 191
4.5. La catégorisation des sons de la langue et de la musique 191
4.6. Quelques réflexions sur la diversité des écoutes musicales 195
5. Conclusions 198
6. Documentation sonore 201
7. Les sons du chapitre 4 202
7.1. Expériences d'écoute 202
7.2. Formes sonores (sources) 202
7.3. Formes sonores : séquences 203
7.4. Formes et flux sonores : groupements et fissions 203
7.5. Ségrégations (spectrale, spatiale) 204
7.6. Anamorphoses sonores 205
7.7. Catégorisation (voyelle, timbre) 205
7.8. Formes et musiques 205

CHAPITRE 5 – **PERCEPTION DES QUALITÉS SONORES : L'INTENSITÉ** 207
1. Introduction 207
2. L'intensité perçue et la zone de sensibilité de l'oreille humaine 208
2.1. Le rôle de la tessiture 208
2.2. Les formants présents dans la zone 3 000 Hz 209
3. La perception de l'intensité et l'enrichissement spectral 210
3.1. Le crescendo musical 210
3.2. Les sourdines, l'effet d'éloignement et d'écho 213
3.3. Le problème des sons graves 213
4. La dynamique des instruments de l'orchestre 214
4.1. La dynamique globale 214
4.2. Les variations de dynamique par la registration 216
5. Jouer fort, jouer doux : le niveau sonore n'est pas seul en cause 217
5.1. Les transformations temporelles et spectrales 217
5.2. L'estimation de l'intensité et la connaissance de la source 218
5.3. La perception d'intensité en contexte temporel et le rôle des silences 220
6. Les nuances d'intensité en musique 221
7. L'émergence par contraste de forme 223
7.1. Les signaux d'avertissement 223
7.2. La notion d'émergence en musique 224

8. Conclusion 226
9. Les sons du chapitre 5 227
9.1. Intensité perçue et spectre 227
9.2. Crescendo musical et changement spectral 227

CHAPITRE 6 – **PERCEPTION DES QUALITÉS SONORES : LA HAUTEUR DES SONS ISOLÉS** 229

1. Préliminaires 229
1.1. Le paramètre « hauteur » dans la musique occidentale 229
1.2. La perception de la hauteur sonore : trois expériences introductives 230
1.3. Bilan des trois expériences : la perception de hauteur et la structure acoustique des sons 233
1.4. La dimension cognitive de la hauteur tonale : expérience de comparaison voix-sifflet 235
2. La hauteur des sons périodiques 237
2.1. Introduction : la singularité des sons purs 237
2.2. Du grave à l'aigu : les bornes de la perception de la hauteur tonale et la tessiture des instruments de musique 237
2.3. La hauteur tonale et le filtrage du fondamental 238
2.4. Les liens entre le contenu spectral et la hauteur tonale perçue 239
2.5. La hauteur tonale et la hauteur spectrale : le contrebasson 240
2.6. Les conflits entre hauteur tonale et hauteur spectrale : quelques sons paradoxaux 242
2.7. La voix humaine : de la mélodie spectrale au chant harmonique 244
3. La hauteur des sons apériodiques 251
3.1. Introduction 251
3.2. Un cas particulier : la quasi-périodicité des sons de piano 252
3.3. La hauteur due à une composante dominante : le diapason à fourche et le marimba 254
3.4. Le modèle harmonique : cloches et timbale 255
3.5. La hauteur des sons apériodiques successifs 260
4. Les sons périodiques modulés : vibratos et trilles 265
4.1. Introduction : les instabilités de fréquence des sons réels 265
4.2. Le vibrato musical : une modulation complexe 266
4.3. Le vibrato de fréquence et la perception de la hauteur 268
4.4. Du vibrato au trille : le rôle du contexte musical 272
4.5. À propos des ornements 274
5. Bilan 276
5.1. La hauteur spectrale et la hauteur tonale 276
5.2. La perception de la hauteur, la facture instrumentale et les modes vibratoires 278
5.3. Perception de la hauteur et mesure acoustique 279
5.4. La hauteur comme qualité de sons connus et catégorisés 279
6. Les sons du chapitre 6 280
6.1. Trois expériences introductives 280
6.2. Sons périodiques 281
6.3. Sons apériodiques 283
6.4. Instabilités, vibrato, trille 285

CHAPITRE 7 – **LA QUESTION DU TIMBRE** 287
1. La musique et le matériau sonore 287
1.1. Le timbre, la musique et l'acoustique 287
1.2. Les définitions acoustiques du timbre : le timbre et son double 288
1.3. Les deux écoutes du timbre : le timbre identitaire et le timbre qualitatif 290
1.4. Le timbre, la musique et ses instruments 291
1.5. Les nouvelles écritures orchestrales 292
1.6. Les nouvelles machines sonores 293
1.7. Les écoutes du matériau sonore : Pierre Schaeffer et les musiques électroacoustiques 294
1.8. La synthèse sonore, pierre de touche de la perception 295
1.9. Du timbre au « son pour lui-même » : vers une dissolution du concept de timbre ? 296
2. Le timbre identitaire des sources instrumentales 297
2.1. La typologie acoustique des formes sonores de base 297
2.2. Typologie des sources sonores impulsionnelles (percussions) 300
2.3. Les sources sonores de hauteur définie 308
2.4. Du son isolé à l'instrument : la variabilité de la forme identitaire avec la tessiture 322
2.5. L'instrument de musique : permanence-variation 332
2.6. Le timbre identitaire et la musique : cohérence et incohérence du timbre 338
2.7. Un bilan : typologie instrumentale et timbre causal 342
3. Le timbre et la catégorisation perceptive 346
3.1. Les catégories instrumentales du timbre identitaire 346
3.2. Caractériser le timbre qualitatif : les mots du timbre 351
3.3. Les attributs perceptifs du timbre : des sons synthétiques aux sons instrumentaux 356
4. Le timbre et les pratiques du son 366
4.1. Les qualités du son : une écoute choisie 366
4.2. La diversité des écoutes dans les pratiques de la qualité sonore 367
4.3. Une méthode d'étude en sémioacoustique 375
4.4. Les problèmes posés par l'évaluation qualitative des instruments de musique et de la voix 379
4.5. Les étapes de la construction acoustique du timbre 382
5. Conclusions 383
6. Les sons du chapitre 7 386

CHAPITRE 8 – **SYSTÈMES D'INTERVALLES ET ACCORDAGE** 391
1. La notion d'intervalle 391
1.1. Données perceptives 391
1.2. Données musicologiques 392
1.3. Les échelles mélodiques de sélection d'harmoniques : le problème de l'harmonique 7 398
1.4. L'estimation d'un intervalle par l'harmonique commun à deux sons 401
1.5. Entre mesure et perception : les intervalles mélodiques et la musique 408
1.6. En conclusion 410
2. Sons stables simultanés : phénomènes physiques 411
2.1. Les battements entre deux sons voisins de l'unisson 411

2.2. Les battements d'intervalles quelconques 413
2.3. Des battements aux sons différentiels : « les sons ont une ombre » 415
2.4. La production de hauteurs complexes avec des instruments à sons entretenus 417
2.5. Les hauteurs d'un son complexe : des notes ? un accord ? un timbre ? 420
3. L'accordage des instruments polyphoniques de hauteur fixe 420
3.1. Les bases acoustiques de l'accordage 421
3.2. La réalisation pratique des systèmes d'accord d'instruments de hauteur fixe 425
3.3. Accordage et instrument : de la théorie à la réalisation pratique 432
4. Le diapason et l'oreille dite absolue 435
5. Les sons du chapitre 8 439
5.1. Les intervalles entre sons successifs (mélodie) 439
5.2. Les intervalles entre sons simultanés 439
5.3. Accordage des instruments polyphoniques à sons fixes 440

CHAPITRE 9 – **VOIX ET PERCEPTION** 443
1. La voix humaine : un instrument très particulier 443
1.1. Une source acoustique polymorphe 443
1.2. Données élémentaires sur l'instrument vocal 445
1.3. La parole et les articulateurs 451
1.4. Les cavités de résonance : voyelles et timbre 451
2. La voix chantée et les cavités de résonance 456
2.1. Interactions d'un son harmonique avec un résonateur 456
2.2. Relations entre la tessiture des voix chantées et les zones spectrales des formants vocaliques 457
2.3. Voyelles et chant lyrique 458
3. Voyelles et chants harmoniques : formants F1 et F2 462
3.1. Les mélodies harmoniques du chant diphonique 462
3.2. Les rapports de fréquence entre formants et fondamentale laryngée 463
3.3. Chant de F1 et technique vocale du period-doubling 464
3.4. Exemples de musique harmonique avec accord d'octave entre F1 et F2 465
3.5. La quintina des chanteurs sardes : une expérience perceptive étonnante 467
3.6. Conclusion 470
4. Les sons du chapitre 9 471

ANNEXES

ANNEXE A – **CONVENTIONS DE NOTATION MUSICALE** 477

ANNEXE B – **INTERVALLES MUSICAUX** 479
1. Tableau des intervalles et de leurs mesures 479
2. Calcul des commas 481
2.1. Calcul du comma syntonique 481
2.2. Calcul du comma pythagoricien 481
2.3. Calcul du comma enharmonique 481

3. Divers 482
3.1. Mesure de la fréquence d'un son avec un accordeur 482
3.2. Trouver la fréquence d'un son dont on connaît l'intervalle par rapport à un autre 482
4. Fréquences des notes du tempérament égal calculées pour l'octave 3 483

ANNEXE C – **PRATIQUE DU LECTEUR MUSICIEN** 485
1. Fréquence, période, célérité, longueur d'onde 485
2. Correspondances note-fréquence-période-longueur d'onde 485
3. Série harmonique 486
4. Un exemple pratique : calculer la fréquence de résonance d'une bouteille 486

ANNEXE D – **TEXTES** 487
1. Analyse de l'écoute d'une mélodie par Edmund Husserl 487
2. L'illusion perceptive du trille 487

ANNEXE E – **VISUALISATION DES PHÉNOMÈNES VIBRATOIRES** 489
1. À propos des animations visibles en ligne 489
1.1. sDR : site de Dan Russel 489
1.2. sPF : site de Paul Falstad 490
1.3. sJW : site de Joe Wolfe 490
1.4. sWR : site de Wolfgang et Rousseau 491
2. Fibrostroboscopie des cordes vocales pendant le chant 491

ANNEXE F – **BIBLIOGRAPHIE** 493
1. Matériel audiovisuel 493
2. Bases de données 493
3. Livres – Périodiques – Articles – Comptes rendus de congrès – Thèses – Rapports 494
4. Internet : sites et pages personnelles 512
5. Logiciels 513

ANNEXE G – **GLOSSAIRE** 515

ANNEXE H – **CONTENU DU DVD-ROM D'ACCOMPAGNEMENT** 531

INDEX 535
INDEX DES NOMS PROPRES 539

INTRODUCTION AU MONDE DES SONS

S'il vous plaît... dessine-moi un son !

Nous n'avons jamais été plongés dans une telle profusion sonore. Pourtant le monde des sons demeure, pour la plupart des auditeurs, un monde mystérieux. Nous ne pouvons ni voir ni toucher ces ondes sonores qui s'entrecroisent et nous enveloppent. Même prisonniers des galettes de cire et des CD numériques, les sons échappent à l'observation courante, car leur nature vibratoire est infinitésimale. La durée d'une oscillation est trop brève pour être saisissable à l'œil nu (quelques millisecondes) et l'amplitude vibratoire du mouvement qui l'a produit extrêmement faible (de l'ordre du micron).

Que retient-on d'un son que l'on vient d'entendre ? Ce que l'on a *compris* d'une parole, ce qu'on a *reconnu* d'une musique, mais pas le son lui-même. Il faut le capter, le fixer, l'écouter et le réécouter, analyser les vibrations pour pénétrer dans la richesse et la complexité de ce que nous entendons quotidiennement et que nous croyons si bien connaître. Lorsqu'enfin nous accédons à une transposition visuelle, les images saisies sur l'écran nous étonnent car elles posent souvent plus de questions qu'elles n'en résolvent. Où sont les *notes* que l'on entend si clairement ? Pourquoi la même mélodie donne-t-elle des images si différentes lorsqu'on change d'instrument ? En quoi consiste le timbre du violon qui nous est si familier et en quoi diffère-t-il de celui de la flûte ? Bien d'autres questions surgissent sur la justesse de jeu, sur les qualités comparées des sons et en particulier sur l'incidence de la salle d'écoute.

Sans prétendre répondre à toutes les questions que se posent les auditeurs, nous proposons d'offrir quelques clés pour entrer dans le monde de l'acoustique par le biais de l'écoute, pour nous approprier l'usage des représentations visuelles du son et, finalement, associer l'écoute à l'analyse acoustique, comme on le fait de la musique et de la parole avec leurs transcriptions écrites, en *dessinant les sons*.

Les musiciens sont les grands magiciens du son, qui est pour eux à la fois le déclencheur et le fil conducteur d'un riche imaginaire personnel. Ils passent de patientes années à maîtriser leur instrument pour produire les sons qui nous touchent. Pour entrer dans l'univers de l'acoustique musicale, ils ont sur les scientifiques l'énorme avantage d'avoir développé à la fois une bonne oreille, une bonne mémoire auditive et une grande mobilité dans les stratégies d'écoute. Ces compétences hautement spécifiques ne doivent pas éclipser celles que chacun de nous possède sans en avoir conscience –- au premier plan desquelles se trouve l'écoute de la parole. Il faut aussi plusieurs années pour apprendre à repérer, dans le flot continu de la parole, les éléments signifiants du langage dont la forme acoustique varie sans cesse, et particulièrement d'un locuteur à l'autre. Comme pour la musique, il s'agit d'une pratique sonore dans laquelle nous sommes à la fois auditeurs et acteurs. Enfin, il existe un domaine sonore, celui de l'environnement, dans lequel nous faisons également preuve d'une compétence auditive insoupçonnée puisque chacun de nous est capable de reconnaître sans effort les sons qui lui parviennent quotidien-

nement, que ce soit dans la rue, dans la campagne ou dans un bistrot. Certes, la musique, la parole et les sons de l'environnement sollicitent différemment nos capacités d'écoute[1], mais, du point de vue acoustique, ce sont des « signaux sonores » qui partagent les mêmes principes de production. Voici donc en guise d'introduction une petite séquence sonore à écouter, ainsi que les « images » des sons entendus sur lesquelles nous reviendrons en détail.

Son 1 (17'')

Consigne d'écoute

Pour écouter le son 1, cliquez sur la barre de lecture du son dans le « livret-sons » (fichier ePub) du présent chapitre ou sur le picto du son dans la marge de la version PDF de l'ouvrage (voir annexe H, page 531, Contenu du DVD-Rom).

Cette séquence offre un concentré d'événements sonores reconnaissables dès la première audition : nous pouvons même imaginer la scène. Dans la rue, une personne répond à un appel téléphonique en élevant la voix car l'environnement dans lequel elle se trouve est particulièrement bruyant. On y perçoit des bruits de moteurs, des klaxons, des cris d'oiseaux, le passage d'une voiture de police, et pour couronner le tout un marteau-piqueur qui se met en route. Comment rendre compte visuellement des événements que nous avons tous repéré avec aisance ?

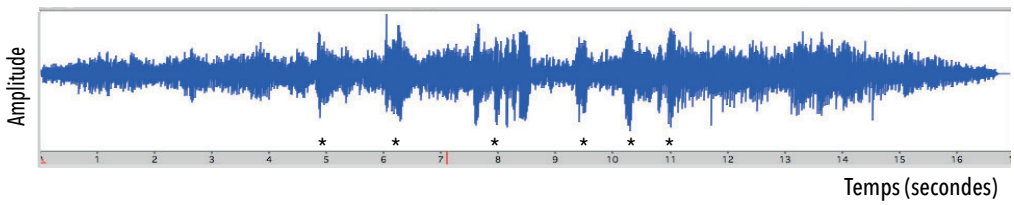


Figure 1 Tracé des variations de l'amplitude sonore en fonction du temps. Les renforcements visibles aux secondes 5 ; 6 ; 8 ; 9,5 ; 10,3 et 11, correspondent aux interventions de la voix qui est au premier plan sonore.

Le tracé de la figure 1, qui représente l'évolution de l'amplitude sonore en fonction du temps, est le plus simple à obtenir. Il convient à l'analyse de sons isolés mais ne permet pas de séparer les événements que nous entendons clairement, car les amplitudes des sons simultanés s'additionnent, notamment les vibrations du bruit de fond urbain noyant celles des autres événements. Un grossissement de l'échelle des temps (zoom) serait de peu de secours.

La représentation de type *sonagramme*[2], apparue dans les années 1950, est une étape majeure de la représentation visuelle des sons, en particulier de ceux de notre environnement quotidien. Il devient possible de différencier les événements selon les zones de fréquence et d'en figurer les variations dans le temps. Sur la figure 2, le degré de noircissement indique la plus ou moins grande intensité du son. Les cris d'oiseaux aigus apparaissent sur la partie supérieure alors que les sons graves des moteurs de voiture et de mobylette occupent la partie inférieure. On repère très bien la structure rythmique régulière de la sonnerie de téléphone ainsi que les fines hachures du marteau-piqueur. Le signal de police apparaît sous forme de raies horizontales disposées en colonnes ; la parole se présente sous forme de courbes ondulantes très variables, en correspondance avec les maxima de la courbe d'amplitude.

1. Pour plus d'informations se reporter au chapitre 4.
2. Nom commercial (francisé) d'une représentation apparue en 1946 aux États-Unis (voir Koenig, W., Dunn, H. K., & Lacy, L. Y., 1946, *The sound Spectrograph*, J.A.S.A., 18 (1), p. 19-49).

Cependant, ni les mots prononcés ni la mélodie typique de la voiture de police ne sont lisibles à première vue.

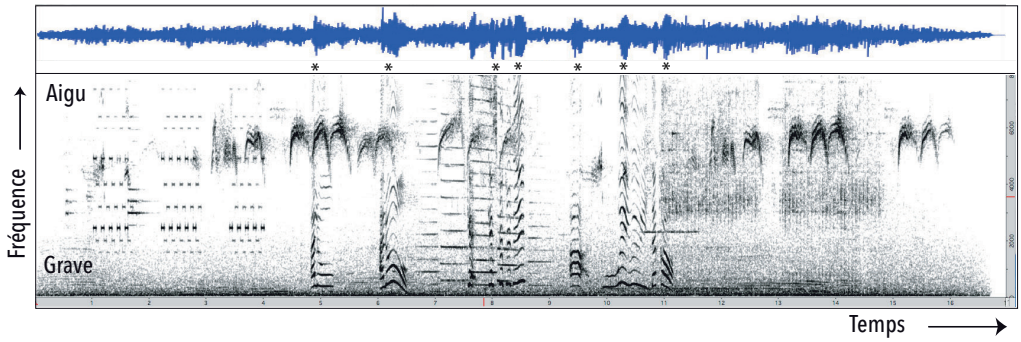


Figure 2 Représentation sonagraphique d'une séquence d'événements sonores captée dans l'environnement urbain.

Cette représentation a encore plus de force lorsqu'il est possible d'associer l'écoute et le déroulement temporel de l'analyse comme dans la séquence vidéo ci-dessous, incluse dans la version PDF de l'ouvrage fournie sur le DVD-Rom d'accompagnement : pour la lire, cliquez sur l'image. Là, la gamme de couleurs évolue des sons les plus faibles (noir, bleu) aux sons les plus forts (jaune, rouge).

Son 1 (17")

Vidéo

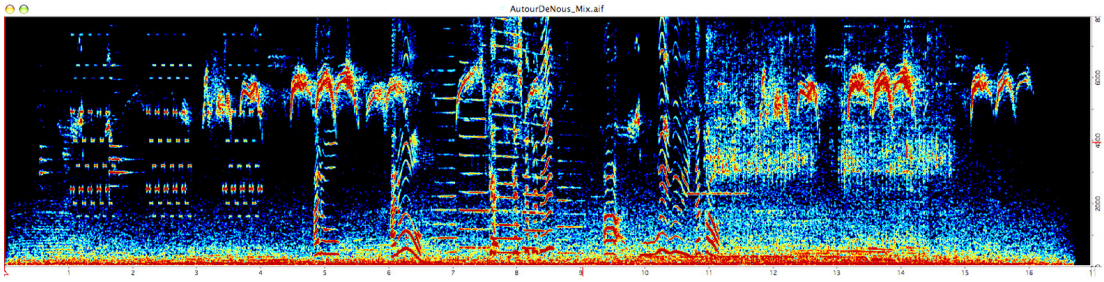


Figure 3 Séquence vidéo captée à partir de la lecture sur écran de l'analyse sonagraphique du Son 1. La gamme de couleurs de l'image est liée à l'intensité sonore. Le rouge figure les sons les plus intenses, le bleu pâle les sons les plus faibles. (Logiciel Audiosculpt)

Cette séquence sonore urbaine est en réalité le résultat d'un mixage des sons enregistrés séparément[3]. Il est donc possible de les écouter individuellement et d'effectuer les analyses des différents types de sons isolément. Le sonagramme de la figure 5 a ensuite été recomposé par la combinaison de six calques auxquels ont été attribués des couleurs arbitraires.

3. Séquence sonore réalisée par Corsin Vogel ; recomposition des calques colorés sur une idée de Charles Besnainou.

Son 2 (3″) | **Son 3 (4″)** | **Son 4 (4″)** | **Son 5 (3″)** | **Son 6 (4″)** | **Son 7 (4″)**

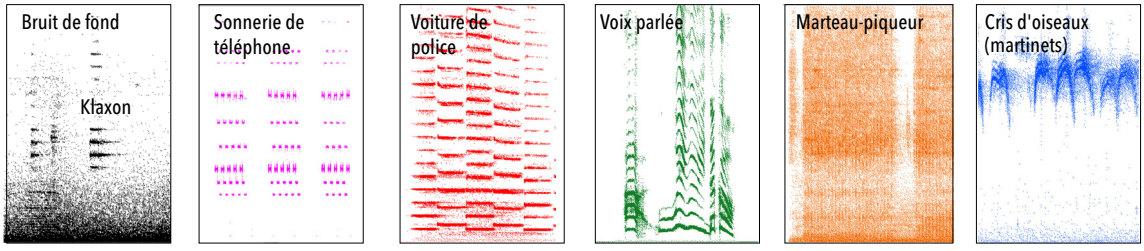


Figure 4 Analyses séparées des différents types sons. Dans le fichier PDF, cliquez sur chaque image pour entendre l'extrait sonore correspondant.

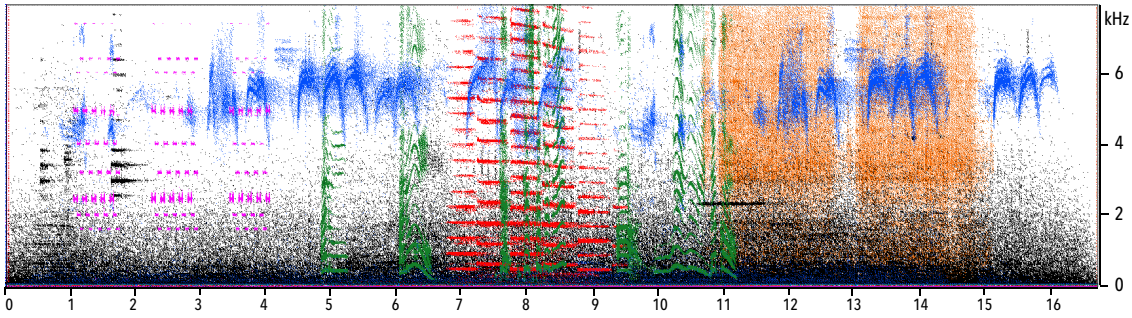


Figure 5 Représentation de type sonagraphique obtenue par la superposition des calques des six séquences de sons analysées séparément. La couleur de l'image est attribuée arbitrairement à un type de son (exemple : le vert pour la parole, le bleu pour la séquence d'oiseaux).

Son 1 (17″)

Séquence urbaine globale

Il est tout à fait remarquable que chaque type de source sonore corresponde à une forme caractéristique reconnaissable visuellement. Les recherches cognitives actuelles conduisent à penser qu'un traitement similaire s'opérerait dans notre système auditif, mais l'appréhension et le traitement de ces « formes sonores » diffère selon les situations d'écoute.

Nous souhaitons attirer l'attention du lecteur-auditeur sur les trois principales situations d'écoute (écoute des sons de l'environnement, d'une conversation ou de la musique) qu'il ne faudra pas perdre de vue par la suite, car les stratégies d'analyses à mettre en œuvre et leurs interprétations en dépendent.

L'écoute des sons de notre environnement – faculté que nous partageons avec les animaux – est celle qui nous permet de repérer et de catégoriser les sources sonores sur une base acoustique, tout en les associant aux événements vécus.

Les sons de l'environnement : quoi ? Où ?

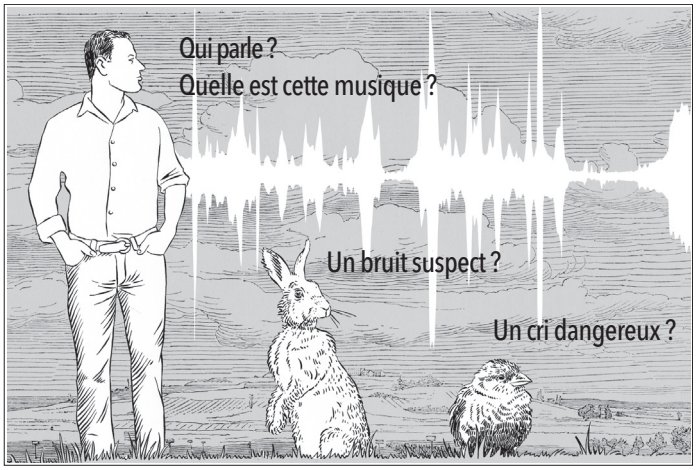


Figure 6 Écoute des sons de l'environnement : décryptage des événements sonores qui se produisent dans le monde extérieur.

Figure 7 Écoute en situation de communication : la conversation. Écoute et « décodage » du sens porté par les sons.

La communication sonore : la parole et la musique

Il en va différemment pour les productions sonores propres aux sociétés humaines que sont la parole et la musique. Elles ont pour support des sources sonores connues, en l'occurrence la voix humaine et les instruments dits de musique.

L'écoute ne s'oriente donc plus sur l'origine des sons mais sur la façon dont ceux-ci sont organisés[4] et sur les qualités qui les caractérisent. Les traitements cognitifs diffèrent aussi selon qu'il s'agit des sons de la langue ou de ceux de la musique.

Toutefois, le substrat sonore qui sollicite l'oreille est commun aux trois situations d'écoute : c'est un signal acoustique dont la production et le comportement physique relèvent des lois de la mécanique. C'est pourquoi le chapitre 1 fournit un rappel des données utiles pour établir les relations existant entre la structure matérielle des corps et la façon dont ils sont mis en vibration, et le chapitre 2 présente les différentes représentations des sons en vue de repérer leur forme acoustique sur les analyses. De façon similaire, le chapitre 3 introduit au fonctionnement de l'oreille humaine. Sur la base de ces éléments, nous pouvons présenter dans le chapitre 4 une approche synthétique de la perception sonore fondée prioritairement sur les propriétés physiques des sons et sur les données de la psychoacoustique, mais en l'inscrivant dans une pratique sonore située, contextualisée, de sorte que les signaux sonores prennent sens pour un auditeur acteur de son écoute.

4. En écho à la définition de John Blacking : « La musique est du son humainement organisé. »

Nous pouvons alors aborder l'étude complexe des modalités d'écoute qui jouent un rôle majeur dans nombre de musiques : l'intensité (chapitre 5), la hauteur (chapitres 6 et 8) et le timbre (chapitre 7). En dernier lieu, un bref chapitre 9 présente quelques productions remarquables de la voix humaine qui combine de façon subtile la quasi-totalité de nos capacités d'écoute.

Figure 8 Écoute de la musique.

Source : Lithographie de Kriehuber, 1846, représentant Franz Liszt au piano entouré de Berlioz et Czerny (debout de gauche à droite), Kriehuber assis à gauche, Ernst à droite.

© Bnf, Paris.

L'abord de cet ouvrage ne nécessite pas de connaissances particulières en acoustique. Il s'adresse à toute personne curieuse du monde des sons et désireuse de mieux connaître la diversité des manières d'entendre. À cet effet nous offrons une grande quantité d'échantillons sonores à expérimenter individuellement, car ils sont le point de départ de notre réflexion. L'auditeur-lecteur devra surmonter de son mieux le divorce des supports matériels de l'écrit et du sonore, et s'imposer, pour une fois, de donner la priorité à l'audition sur la vision.

Les sons de l'introduction

Son 1 – Séquence sonore rassemblant plusieurs événements reconnaissables : bruit de fond de circulation ; parole, sonnerie de téléphone ; signal de police ; oiseaux (martinets) ; marteau piqueur. Mixage : C. Vogel, M. Castellengo. Idée du montage calque : Ch. Besnainou. [Archives LAM]

Son 2 – Bruit de fond de circulation avec klaxon.

Son 3 – Sonnerie de téléphone.

Son 4 – Signal de la voiture de police.

Son 5 – Voix de la conversation au téléphone.

Son 6 – Marteau piqueur.

Son 7 – Cris d'oiseaux en vol (martinets noirs).

CHAPITRE 1

DES VIBRATIONS AUX SONS DE LA MUSIQUE

Le silence éternel de ces espaces infinis m'effraie. Pensées, Blaise Pascal

1. À l'origine du son : le mouvement

Pour commencer nous évoquerons l'impression étrange que l'on ressent lorsque l'on entre pour la première fois dans une pièce anéchoïque, plus couramment appelée « chambre sourde ». Dans ce lieu isolé des bruits environnants et garni de matériau absorbant règne un silence qui surprend. Les seuls sons que nous percevons, qui prennent une dimension nouvelle, presque angoissante, proviennent du fonctionnement interne de notre propre corps : battements du cœur, sifflements d'oreille. Nous prenons soudain conscience du lien direct entre son et mouvement. En effet, la chambre sourde est un lieu spécialement construit pour y effectuer des mesures acoustiques dans le plus grand silence. C'est donc un lieu privé de vie, isolé du monde extérieur, au sein duquel aucun événement accidentel ne peut se produire. En sortant, nous retrouvons avec plaisir le bruit des pas amplifié par la résonance du couloir, puis le brouhaha extérieur, signe sonore de la présence active des êtres qui sont autour de nous. Le monde dans lequel nous vivons *bruit* des mouvements naturels du vent et de l'eau et de ceux des êtres vivants qui se meuvent, communiquent ou produisent de la musique pour leur seul plaisir.

À l'origine d'un son, il y a donc un mouvement et, pour effectuer un mouvement, il faut dépenser de l'énergie. En d'autres termes, lorsqu'on entend un son, il s'est produit – ou il est en train de se produire – un événement, proche ou distant, dont l'onde sonore porte la trace matérielle jusqu'à nos oreilles.

2. La production et la propagation des sons

2.1. En bref

Depuis le mouvement initial jusqu'au son entendu, plusieurs processus sont mis en œuvre (figure 1.1).

- L'*excitation,* qui est à l'origine du mouvement, apporte plus ou moins d'énergie, ce qui produit un son plus ou moins intense. Si elle est de nature impulsionnelle – un choc, un pincement comme pour la guitare –, la totalité de l'énergie est fournie au moment de l'impulsion. Si elle est de nature entretenue – frottement, souffle comme pour la flûte –, l'énergie est apportée tout au long

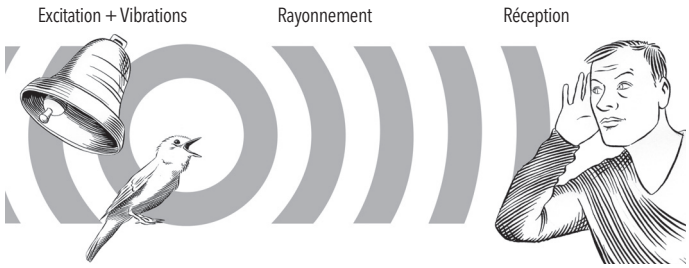


Figure 1.1 De l'excitation à la réception d'un son.

de l'excitation. Cette distinction est capitale, car elle conditionne le contenu physique du son : son évolution temporelle et son effet perceptif.

- La *vibration* peut prendre des formes diverses selon le type de corps excité : cavité aérienne, corde vibrante, plaque, caisse. Dans la réalité, le corps excité est toujours un système complexe subissant plusieurs sortes de vibrations simultanées : les modes vibratoires. Dès le premier instant de la vibration, des ondes se propagent de proche en proche, depuis le point d'excitation jusqu'à la totalité de la structure vibrante. Plusieurs sortes d'ondes (compression, cisaillement, déplacement) cœxistent au sein de la structure vibrante.
- Le *rayonnement*. Aux interfaces des surfaces vibrantes (table d'harmonie) et de l'air environnant, ainsi qu'aux orifices des cavités (pavillon, trous latéraux des instruments à vent), naissent des ondes aériennes qui assurent la propagation du son depuis la structure vibrante jusqu'à un récepteur (oreille, microphone). L'air environnant est traversé d'ondes dont les amplitudes s'additionnent et qui forment un champ d'interférences complexe.
- La *réception* ou saisie du son. Le point où l'on place le microphone est toujours un point particulier du champ sonore. Dans la situation ordinaire d'un auditeur, chaque oreille capte un signal différent. Ajoutons qu'une partie des vibrations peut aussi être transmise directement par voie solidienne[1].

Une petite synthèse est proposée dans le tableau 1.

Tableau 1. Production des sons : étapes et processus

EXCITATION	VIBRATION	RAYONNEMENT	RÉCEPTION DU SON
Apport d'énergie : impulsion ou entretien.	Ébranlement d'une structure matérielle déformable et génération d'ondes dans la structure. Modes vibratoires à 1 ou 2 dimensions.	Propagation d'ondes dans le milieu environnant (gazeux, liquide, solide).	Capteurs de vibrations : oreille humaine ou capteurs mécaniques (microphones, accéléromètres).

Nous ne développerons que les aspects de la production des sons qui sont nécessaires à la compréhension de la perception sonore, et plus particulièrement ceux de la hauteur et du timbre. La présentation qui suit est donc volontairement succincte, et nous engageons le lecteur à consulter le glossaire, annexe G (pages 515 à 536) où sont développés quelques compléments et définitions utiles. Enfin, nous

1. Citons le violoniste dont le menton est en contact direct avec la caisse de l'instrument et surtout le chanteur dont les organes du larynx vibrent avec intensité, non loin de l'oreille.

utiliserons dès maintenant la représentation spectrographique des sons, nous réservant d'entrer dans le détail des différentes techniques d'analyse du son et du choix des grandeurs à représenter dans le chapitre suivant.

2.2. L'excitation : impulsion ou entretien

2.2.1. Les deux modalités de fourniture d'énergie

La première étape de la production sonore est l'apport d'énergie nécessaire pour produire une vibration. La quantité d'énergie conditionne l'amplitude du mouvement et, de façon plus ou moins directe, l'intensité sonore. La façon la plus simple de mettre un corps en mouvement est de lui donner un choc. Lorsqu'on frappe une cloche, une corde, la membrane d'un tambour, toute l'énergie est fournie dès le départ au moment de l'impulsion initiale. La vibration s'arrête lorsque l'énergie est dissipée. On dit que l'excitation est de type impulsionnelle. L'impact déforme une zone de la structure matérielle qui tend à reprendre sa position initiale : c'est l'origine de la vibration. Il faut cependant que la matière ainsi déformée soit élastique (voir Glossaire, page 518) et ne se comporte pas comme de la pâte à modeler, dont on sait qu'elle ne « sonne » pas, car la déformation s'y imprime en permanence, sans produire de réaction en retour. Les vibrations des corps solides (métal, bois, verre, corde tendue) sont sensibles au toucher et parfois même visibles. Il n'en va pas de même pour l'air, bien que ce milieu matériel peu dense soit également susceptible de vibrer. On peut exciter l'air contenu dans un tuyau ou dans une cavité en produisant une impulsion ou une variation brusque de pression (positive ou négative) à l'une des ouvertures du tube. C'est le « pop » de la bouteille qu'on débouche, le *tongue ram* du flûtiste[2] ou encore les bruits de clés des instruments à vent. Cependant, les vibrations aériennes dues à une seule impulsion s'arrêtent très vite. Aussi, l'excitation habituelle des instruments à vent est-elle plutôt de type « entretenu », c'est-à-dire que l'énergie nécessaire à la vibration est fournie continûment par le souffle du musicien. Les autres instruments entretenus utilisent principalement le frottement d'un archet (violon), du doigt (harmonica de verre, cristal Baschet) ou encore d'une pièce de bois (bol tibétain)[3].

L'énergie dont dispose un être humain est nécessairement limitée en quantité et en durée. Étant donné que notre oreille est moins sensible aux sons graves (50 Hz) qu'aux sons aigus (3000 Hz), la plupart des sources sonores mécaniques, comme les sifflets, la voix et les instruments de musique traditionnels, se sont adaptées aux meilleures zones de réception de l'oreille. Nous verrons que ces sources nous font entendre indirectement les sons graves au travers des harmoniques élevés (voir chapitres 3 et 6). C'est aussi la raison pour laquelle il n'y a pas de relation simple entre l'énergie fournie, l'amplitude des mouvements vibratoires et l'intensité perçue, car celle-ci dépend de la fréquence des vibrations (voir chapitre 5).

En résumé, lorsque l'apport d'énergie de l'excitation à l'origine du son se produit en une seule fois (la guitare), *le son est de type impulsionnel* ; lorsque l'énergie est fournie continûment (la flûte), *le son est entretenu*. Cette distinction est capitale pour comprendre la structure acoustique des sons et leur perception.

2. Technique de jeu produite avec la langue qui obture violemment le trou d'embouchure.
3. Dans ce cas, le doigt et le bout de bois se comportent exactement comme un archet de violon.

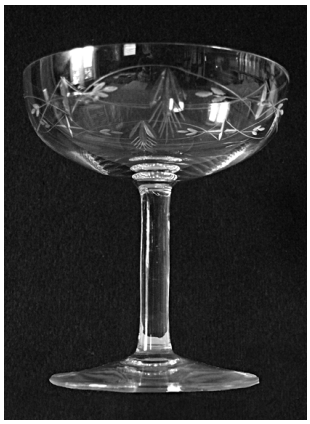


Figure 1.2 Un verre de cristal prêt à chanter.
Cliché M. C.

2.2.2. Un exemple pratique : le verre de cristal

Pour cette expérience, il faut de préférence disposer d'un verre à pied ou d'une coupe dont le bord est fin (figure 1.2). En maintenant par sa base le verre posé sur une table, on le pince[4] brièvement en saisissant le bord avec les ongles. Le verre « sonne » un peu à la façon d'une petite cloche. Chacun sait qu'on peut aussi le faire sonner de façon prolongée en appliquant sur le bord un doigt humide et propre que l'on fait glisser légèrement dans un mouvement circulaire continu. Pour que l'opération réussisse, il faut appuyer convenablement – ni trop, ni trop peu –, c'est-à-dire effleurer de façon sensible le verre. Dès que le verre chante, le doigt se comporte comme un archet qui « accroche » le bord et le met en vibration ; il ne faut donc pas l'étouffer en appuyant trop pesamment.

Les exemples sonores **1.1** et **1.2** donnent à entendre les sons obtenus par ces deux modes d'excitation, qui sont très différents. Sur la figure 1.3, la courbe supérieure représente la variation de l'amplitude en fonction du temps. Le son impulsionnel (verre percuté) a une amplitude maximale dès le début de l'excitation, puis il décroît. Le son entretenu (verre frotté) commence progressivement et persiste tant que le doigt reste en contact avec l'objet. Les variations d'intensité sont dues au déplacement du point d'excitation par rapport au microphone[5].

Son 1.1 (5″)
Verre percuté

Son 1.2 (5″)
Verre frotté

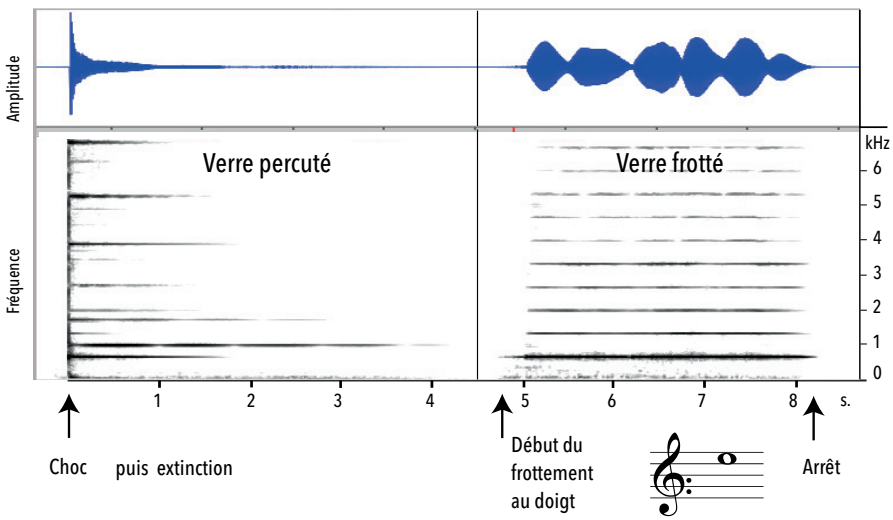


Figure 1.3 Le même corps excité de deux façons différentes : à gauche par un choc, à droite par un frottement. Analyse de l'amplitude (partie supérieure) et du contenu en fréquence (partie inférieure) des deux types de sons.

À l'écoute, le verre ne donne pas du tout la même sensation de hauteur sonore selon le type d'excitation. Sur la partie inférieure de la figure 1.3, chaque raie horizontale correspond à une fréquence émise par le verre. Le verre frotté émet une note bien définie, un mi_4[6], qui apparaît sur la figure sous forme d'un réseau de raies horizontales équidistantes. Le verre percuté produit un son plus complexe à écouter. Il

4. Nous utiliserons indistinctement les termes « pincé », « percuté » ou « tinté » pour signifier que le verre est mis en vibration par un choc léger.
5. En tournant, le doigt change de position et d'orientation par rapport au microphone qui est fixe (voir Glossaire, *Rayonnement* (quadripôle)).
6. Pour les conventions d'écriture et de numérotation, voir l'annexe A.

débute par un choc (trait vertical indiqué sur la figure), puis on entend le mi_4 mais aussi d'autres notes. Les raies horizontales sont irrégulièrement espacées et évoluent pendant l'extinction. La largeur de ces raies diminue avec le temps : dès la troisième seconde ne subsiste plus que la deuxième raie en partant du bas, celle qui fait entendre la quinte.

Ces analyses montrent que ce que nous appelons « un son » est un événement complexe, caractérisé par une courbe dont l'amplitude temporelle est typique du mode d'excitation et par un ensemble de fréquences pouvant donner, selon les cas, la sensation d'un accord complexe ou d'une note unique. Avec un même corps, ici le verre, nous avons produit deux sons très différents. En les écoutant plus attentivement, on peut toutefois y percevoir des similarités : notamment, la note musicale du verre frotté, s'entend nettement au tout début du son du verre percuté.

Son 1.3 (7″)

Vibraphone : percuté puis frotté

2.2.3. Autre exemple : la lame de vibraphone et la corde de harpe

On peut appliquer ces deux modes d'excitations, choc ou frottement, à des corps quelconques. Voici deux autres exemples produits, l'un avec une lame de vibraphone (fa_2), l'autre avec une corde de harpe (mi_2).

Son 1.4 (7″)

Corde de harpe : pincée puis frottée

Bien que l'entretien de la lame de vibraphone par un archet soit très court (300 ms), on retrouve à l'écoute la même opposition de sonorité entre le son percuté, assez complexe, et le son frotté faisant entendre clairement une note (Son **1.3**). En revanche, la corde de harpe pincée ou frottée donne deux sons similaires à l'écoute (Son **1.4**), qui ne diffèrent que par leur allure temporelle, décroissante ou entretenue (figure 1.4).

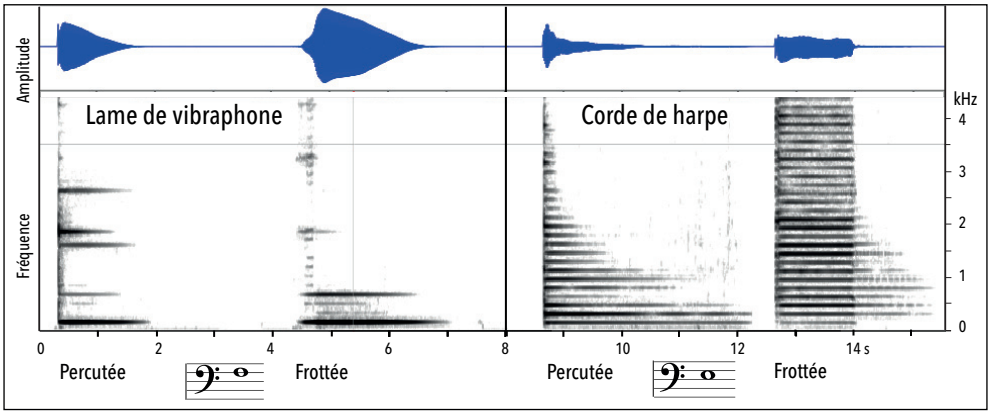


Figure 1.4 Excitation par impulsion ou par frottement appliquée à une lame de vibraphone et à une corde de harpe. Partie supérieure : amplitude/temps. Partie inférieure : spectrogramme temporel.

Les différences constatées entre les sons du verre, de la lame ou de la corde, selon que ces objets sont percutés ou frottés, proviennent principalement de la structure géométrique mise en vibration. Le verre et la lame de vibraphone sont des structures volumineuses dont on doit étudier les déformations dans les trois dimensions de l'espace, alors qu'on peut considérer qu'une corde tendue est déjà bien représentée par sa seule longueur. Nous abordons maintenant la deuxième étape : celle de la vibration des corps matériels.

2.3. La notion d'onde

2.3.1. La propagation des ondes

Il est habituel d'introduire la notion d'onde en prenant pour exemple une étendue d'eau calme à la surface de laquelle on lance un objet (figure 1.5). Cette expérience commune à tous les pêcheurs à la ligne a le grand mérite d'offrir une visualisation claire d'un phénomène oscillant et de sa propagation sous forme d'ondes circulaires. Les cercles grandissants rendent visible la propagation du mouvement sans déplacement de matière : si l'on dispose d'un objet flottant placé non loin du point d'impact, on constate qu'il oscille au passage des cercles tout en restant à la même distance de celui-ci.

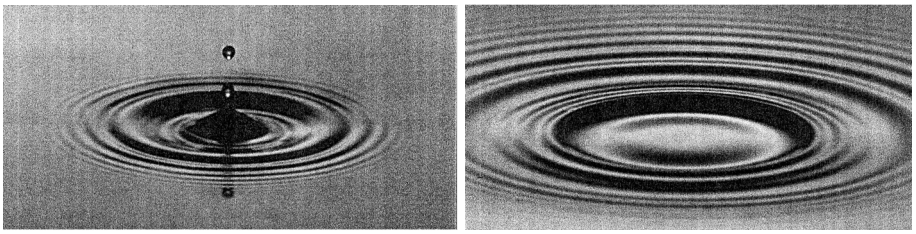


Figure 1.5 Ondes concentriques dues à la propagation d'une oscillation localisée au point d'impact.
DR.

L'expérience que nous venons de décrire met en jeu deux ordres de phénomènes étroitement liés : d'une part un mouvement oscillant localisé, d'autre part la propagation de ce mouvement sous forme d'une onde qui s'étend, de proche en proche, à une zone de plus en plus grande. Au cours de la propagation, l'amplitude des oscillations décroît en fonction de l'éloignement au point d'impact et une partie de l'énergie se dissipe au cours de la transmission. Le mouvement s'amortit et, finalement, il s'arrête lorsque l'énergie totale a été consommée.

La vitesse de propagation (ou célérité) des ondes dépend des caractéristiques mécaniques du milieu. Elle est d'environ 340 m/s dans l'air à 15 °C, ce qui est beaucoup plus faible que pour la lumière. Lors d'un orage, le décalage entre la célérité de l'onde lumineuse de l'éclair (300000 km/s), quasi instantanée pour nous, et le bruit du tonnerre survenant, par exemple, 3 s plus tard, permet d'apprécier la distance de l'orage : environ 1 km dans notre exemple.

2.3.2. Les ondes dans les instruments de musique : exemple du cor des Alpes

Par opposition à une grande étendue d'eau, les parties vibrantes des instruments de musique sont de petites dimensions. Lorsqu'on met en vibration la membrane d'une timbale, une corde de guitare ou l'air contenu dans un tuyau, les oscillations nées au point d'excitation se propagent jusqu'aux bords de la caisse (ou aux extrémités de la corde ou du tuyau) qu'elles atteignent avec une amplitude encore notable. Là, elles se réfléchissent, repartent en sens inverse, et croisent les ondes incidentes en produisant un « champ d'interférences » (voir Glossaire, page 521).

Son 1.5 (2″)

« Pop » : impulsion sur l'embouchure du tuyau

Le cor des Alpes est constitué d'un tuyau conique d'environ 4 m de long (figure 1.6). Plaçons un microphone près du pavillon et, avec la paume de la main, produisons sur l'embouchure une petite surpression qui progresse le long du tuyau jusqu'au pavillon (Son **1.5**). Arrivée au pavillon, une petite partie de la perturbation continue sa progression dans le milieu extérieur et atteint le microphone, l'autre partie se réfléchit et revient vers l'embouchure d'où elle repart pour un nouveau voyage vers le pavillon. Entre deux rencontres avec le microphone, la perturbation parcourt donc le double de la longueur du tuyau, soit environ 8 m.

Figure 1.6 Joueur de cor des Alpes.

Source de l'image : disque Odéon *Ländlerkapelle Balmerbuebe Wilderswil.* MSOE 4054 (1959)

La figure 1.7 montre l'analyse du signal d'amplitude recueilli par le microphone. On voit que l'impulsion initiale effectue plusieurs allers-retours dans le tuyau, avec une amplitude décroissante, avant de s'amortir totalement. Quelle est la durée séparant deux arrivées ? En sélectionnant l'intervalle entre deux impulsions successives, le logiciel d'analyse[7] affiche 0,021s (zone bleue). Cette valeur est proche de celle que l'on peut estimer en calculant la durée de propagation dans le tuyau : longueur du tuyau (m)/célérité du son (m/s), soit 8/340 = 0,023 s. Ce n'est qu'un ordre de grandeur, car ni la longueur effective du tuyau ni la célérité du son ne sont connues précisément.

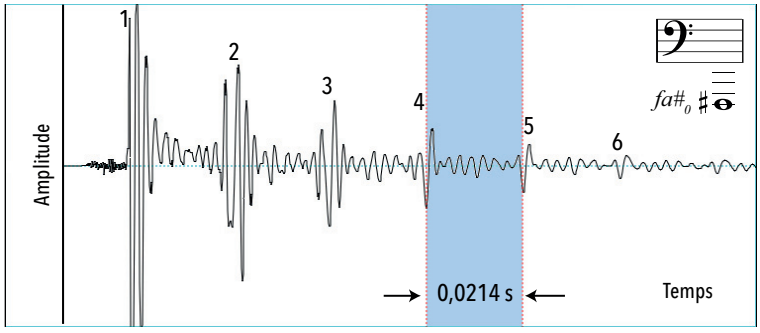


Figure 1.7 Variation de l'amplitude en fonction du temps. Succession des impulsions se propageant dans le tuyau avec une amplitude décroissante, recueillies au pavillon d'un cor des Alpes.

À l'intérieur du tuyau (tube conique de longueur finie), les ondes aériennes restent en partie prisonnières en effectuant des mouvements réguliers d'aller-retour entre les deux extrémités. La petite proportion qui sort du tuyau (entre 5 et 10 % de l'énergie totale, ce qui est très faible) constitue le son rayonné. Le microphone recueille un train d'impulsions régulièrement espacées de 0,0214 s. La fréquence du son est inverse de la durée séparant deux impulsions, soit : 1/0,0214 s = 46,72 Hz. C'est un son grave, voisin d'un $fa\#_0$, difficile à apprécier par l'oreille, car sa durée est brève : l'énergie de l'impulsion initiale se dissipe rapidement au cours de la propagation. Entretenir le son, c'est apporter de l'énergie de façon synchrone à la fréquence du système. Les phénomènes réels sont extrêmement compliqués. Il nous suffira de dire ici que si nous produisons de nouvelles impulsions « en phase »

7. Praat.

avec les ondes de retour, de sorte que leurs amplitudes s'additionnent à chaque cycle, le système accumule de l'énergie. Au bout d'un certain nombre d'allers-retours, s'établit une configuration spatiale appelée « déformée modale ». En certains points, l'air reste immobile – ce sont les « nœuds de vibrations » – alors que d'autres vibrent avec une grande amplitude – ce sont les « ventres de vibration ». L'instrument fonctionne alors en régime entretenu et le tuyau est le siège « d'ondes stationnaires ».

2.4. Les modes vibratoires (transition ondes-modes)

2.4.1. Les fréquences propres

Une déformée modale, comme celle que nous venons de décrire, ne s'établit que pour une fréquence de vibration particulière qui est appelée « fréquence propre ». Déformée modale ou fréquence propre sont deux façons – l'une spatiale, l'autre temporelle – de caractériser cet état particulier d'un système vibrant dans lequel les ondes sont stationnaires. L'ensemble constitue un « mode propre », lequel dépend des caractéristiques géométriques du système et de la vitesse de propagation des ondes dans le milieu[8].

2.4.2. Les modes propres et la géométrie des corps vibrants

La notion de mode propre est fondamentale en acoustique puisqu'elle gouverne tous les systèmes vibrants, y compris les lieux d'écoute. En acoustique musicale, l'ajustement des modes propres guide la conception et la réalisation de la plupart des instruments, qu'il s'agisse des fréquences de jeu (instruments à vents, barres et cordes vibrantes) ou de l'amplitude des composantes spectrales qui contribuent à la qualité du timbre (principalement la caisse des instruments à cordes).

Les modes propres correspondent aux déformations qu'adopte un corps qu'on excite et dépendent donc en premier lieu de la forme géométrique de celui-ci. L'étude théorique distingue trois catégories :

- les corps dont les déformations se produisent principalement le long d'une seule dimension (1D) : les cordes, les tuyaux ;
- ceux dont les déformations affectent les deux dimensions d'une surface (2D) : les membranes, les tables d'harmonie ;
- ceux dont les déformations affectent un volume, une « coque » : les cloches, les caisses de résonance.

Le deuxième paramètre à prendre en compte pour chaque type de corps concerne la façon dont il est construit, en particulier les points de fixation ou d'encastrement (qui sont des zones de déplacement quasi nul, donc des nœuds de vibration), ainsi que les caractéristiques du matériau, notamment son homogénéité. À titre d'exemple, une barre d'aluminium (vibraphone), dont les caractéristiques mécaniques sont les mêmes dans toutes les directions, se comportera différemment d'une barre de bois (marimba), matériau inhomogène, qui ne possède pas les mêmes caractéristiques de propagation des vibrations dans le sens des fibres ou dans le sens perpendiculaire aux fibres.

Quelques représentations animées aident à comprendre les modes vibratoires (voir Annexe E et Glossaire). Celles des structures longilignes comme la corde ou le tuyau – dont on néglige le diamètre – sont les plus faciles à comprendre. Connaissant la longueur, les conditions aux extrémités et la vitesse de propagation des ondes dans une corde ou dans l'air du tuyau, on peut définir les modes propres et calculer les fréquences propres du système. Le problème est plus compliqué pour

8. Cette présentation est volontairement simplifiée. Pour un développement de ces notions voir Chaigne & Kergomard, 2008, chapitre 3, p. 98-100.

les structures 2D. Nœuds et ventres de vibration ne sont plus situés sur un seul axe. La surface vibrante se divise en zones ventrales séparées par des lignes nodales. Pour une même forme, plusieurs familles de modes peuvent se produire. Les structures volumiques nécessitent de considérer les modes s'établissant dans les différents plans de coupe. Enfin, dans les trois catégories de systèmes peuvent se produire des modes vibratoires dus à différentes sortes d'ondes : ondes transversales, ondes longitudinales (ou de compression), ondes de torsion. Il faut toutefois insister sur le fait qu'au moment de l'excitation, *tous* les modes vibratoires coexistent tant bien que mal. Nous ne développerons pas plus et nous engageons les lecteurs curieux à se reporter aux ouvrages cités en bibliographie. Pour une première approche en acoustique, le lecteur pourra consulter : Leipp, 2011 ; Pierce, 1984 ; Rossing, 2002 ; Winckel, 1960, et plusieurs chapitres du *Livre des techniques du son* édité par Mercier (Liénard, Castellengo, Vivié et Cassan). Les traités de Bruneau, Chaigne et Kergomard, Fletcher, ainsi que la collection des ouvrages de Bouasse s'adressent aux lecteurs déjà avertis.

2.4.3. La matérialisation des modes vibratoires des plaques : figures de Chladni

Nous avons dit que tous les corps possédaient des modes propres. Ceux des cordes furent explorés et théorisés très tôt, grâce à la pratique du monocorde, et servirent de modèle à la théorie des tuyaux sonores (voir chapitre 8, § 1.2.3). Les rapports de fréquences remarquables entre les modes successifs trouvèrent leur application dans la théorie musicale, en particulier chez Joseph Sauveur qui développa la théorie des harmoniques du Plein Jeu d'orgue (Sauveur, 1702). Les modes de vibration des plaques ne furent explorés qu'au début du XIXe siècle par un physicien expérimentateur, Ernst Chladni, inventeur du clavicylindre[9]. Chladni eut l'idée de répandre un peu de sable à la surface d'une plaque métallique fixée en son centre et de la faire vibrer au moyen d'un archet[10]. L'expérience, assez aisée à reproduire avec du sel fin, permet de produire des figures géométriques spectaculaires (figure 1.8).

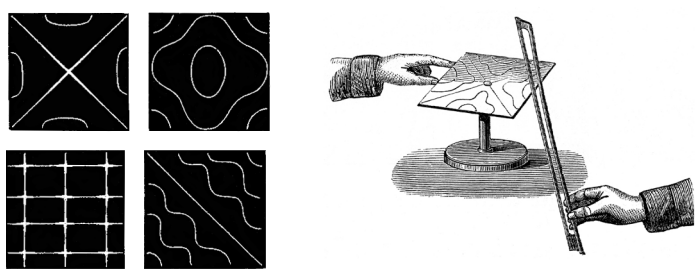


Figure 1.8 Figures de Chladni : matérialisation de la forme vibratoire d'une plaque rectangulaire à une fréquence donnée. À droite, remarquez la pose des doigts sur la plaque pour imposer l'emplacement des nœuds.

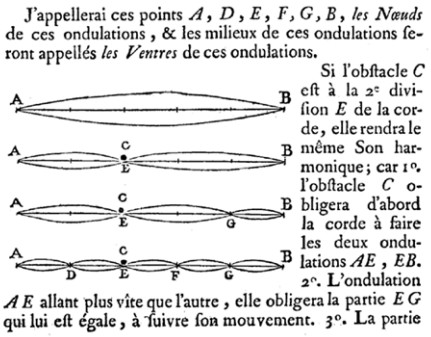
Source : Tyndall, J., 1869, *Le son* (traduction française de l'abbé Moigno). Paris : Gauthier-Villars, figures 60 et 61.

Lorsque l'archet accroche un mode, on entend un son de fréquence stable et le sel se rassemble selon les lignes nodales du système. Pour contraindre la plaque à vibrer sur un mode particulier, il faut poser légèrement les doigts sur le bord, à l'endroit de futures lignes nodales et exciter avec l'archet à l'endroit d'une zone ventrale (voir figure 1.8, à droite).

9. Instrument à tiges de verre qu'il présenta à l'Académie des sciences de Paris en 1808.
10. Dans la préface de son traité d'acoustique, publié en français en 1809, Chladni précise que cette idée lui est venue après avoir vu les figures « électriques » que Lichtenberg avait publiées avant lui en 1777. Voir aussi N. Witkowski, 2001, *Une histoire sentimentale des sciences*, Le Seuil, Paris, p. 135-139.

2.4.4. Les modes propres et les « harmoniques » effleurés

Les instrumentistes ont une connaissance empirique des différents modes propres et de leur sélection. Dans le jeu ordinaire, les violonistes utilisent normalement le premier mode vibratoire de la corde vibrante, mais quelquefois ils jouent des « harmoniques effleurés ». Tout en tirant l'archet, ils posent légèrement le doigt sur la corde, ce qui a pour effet de gêner la formation d'un ventre de vibration ou, ce qui revient au même, d'imposer un nœud de vibration à cet endroit. Les points favorables aux harmoniques effleurés, connus depuis longtemps, sont le milieu, le tiers, le quart de la corde, etc., c'est-à-dire les endroits susceptibles de provoquer la subdivision de la corde en un nombre entier de parties (voir figure 1.9).


J'appellerai ces points *A*, *D*, *E*, *F*, *G*, *B*, *les Nœuds* de ces ondulations, & les milieux de ces ondulations seront appellés *les Ventres* de ces ondulations.

Si l'obſtacle *C* eſt à la 2e diviſion *E* de la corde, elle rendra le même Son harmonique; car 1°. l'obſtacle *C* obligera d'abord la corde à faire les deux ondulations *AE*, *EB*. 2°. L'ondulation *AE* allant plus vîte que l'autre, elle obligera la partie *EG* qui lui eſt égale, à ſuivre ſon mouvement. 3°. La partie

Figure 1.9 Démonstration expérimentale des modes vibratoires d'une corde vibrante.

Sauveur, J., 1701, *Mémoires de l'Académie des Sciences*, p. 352.

De façon analogue, le débouchage d'un trou au milieu du tuyau d'une flûte fait passer au deuxième mode vibratoire, voisin de l'octave du premier. Il existe même des instruments dont les fréquences de jeu sont exclusivement celles des modes propres (voir § 5.2.2, page 25). Précisons dès maintenant que l'emploi du mot « harmonique » dans le contexte musical a un sens très différent de celui que nous lui donnons en acoustique (voir page 24).

2.4.5. Un exemple pratique : les modes vibratoires du diapason à fourche

Le diapason à fourche est un instrument construit pour produire une fréquence étalon, le plus souvent un la_3 dont la fréquence de référence est 440 Hz. Encore faut-il apprendre à « en jouer ».

Son 1.6 (5")

Diapason à fourche

Lorsqu'on frappe le diapason, on entend d'abord un choc métallique suivi d'un son très bref et très aigu (Son **1.6**). Ce son, que nous appellerons « partiel 2 », correspond à la fréquence propre du mode 2 du diapason (figure 1.10). Il s'amortit très rapidement[11]. La fréquence d'accord du la_3 est celle du partiel 1 correspondant au mode 1 du diapason. En pratique, le partiel 2, dont la fréquence est ici environ 2800 Hz, est ignoré perceptivement (figure 1.11).

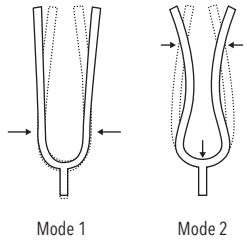


Figure 1.10 Modes vibratoires d'un diapason à fourche.

11. Le diapason se comporte comme une verge encastrée dont la fréquence du deuxième mode vibratoire est environ six fois celle du premier.

Le rayonnement acoustique du diapason tenu par sa tige est très faible : on doit l'approcher de l'oreille pour l'entendre. Or, si l'on met la tige en contact avec une table d'harmonie, les ondes de compression provoquées par les vibrations des branches se transmettent par voie solide à toute la surface. Celle-ci vibre en produisant des ondes aériennes de plus grande amplitude.

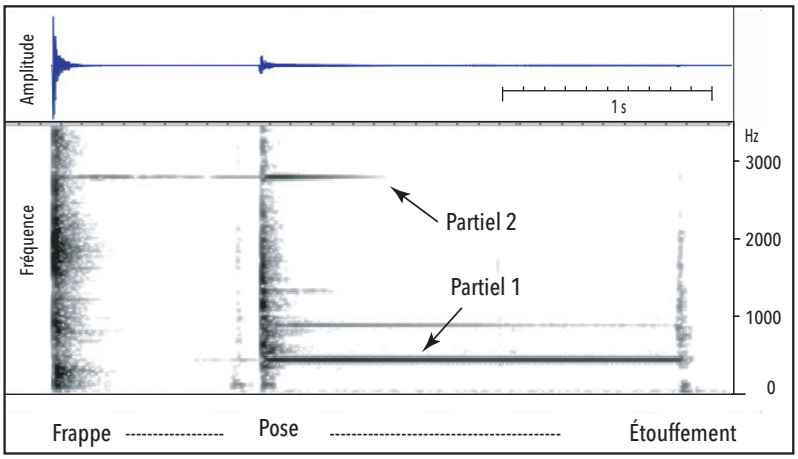


Figure 1.11 Analyse acoustique du son produit. Le mode 1 donne la fréquence d'accord (Partiel 1 sur l'analyse). Le mode 2 produit un son aigu que l'on entend au moment du choc et qui s'évanouit rapidement (Partiel 2 sur l'analyse).

La *vibration* peut prendre des formes diverses selon le type de corps excité. Dans la réalité instrumentale, le corps excité est toujours une structure complexe subissant plusieurs sortes de vibrations simultanées : **les modes vibratoires**.

2.5. Des vibrations de la structure au son rayonné

Lorsque nous entendons le son d'une guitare ou d'une flûte, situés à une certaine distance de nous, cela signifie que les vibrations que produisent ces instruments ont ébranlé l'air environnant qui les a transmises jusqu'à nos oreilles. L'exemple de la flûte est intuitif puisque son fonctionnement a pour fondement des modes vibratoires aériens. Chaque orifice (embouchure, pavillon, trou latéral) se comporte comme une source vibratoire. Pourtant, les ondes rayonnées par les orifices ne sont dues qu'à des « pertes » du système évaluées à environ 5 % de l'énergie totale, laquelle reste confinée dans le tube où elle contribue à l'entretien de la vibration.

Le problème se pose différemment avec les instruments à cordes. Une corde qui vibre ne déplace qu'un très faible volume d'air autour d'elle. Tendue sur une structure rigide et inerte, une corde vibrante est quasiment inaudible. Les cordes sont donc toujours couplées à des surfaces (bois, peau) qu'elles déforment en vibrant et qui, à leur tour, génèrent des ondes de compression aériennes. Étant donné que les vibrations des différentes parties de la caisse, et plus particulièrement celles de la table, jouent un rôle majeur dans le rayonnement du son, il y a donc lieu de prendre en compte pour ces instruments non seulement les modes propres des cordes mais aussi ceux des plaques auxquelles elles sont fixées. Nous avons vu dans

l'expérience précédente du diapason que celui-ci n'est audible à distance que couplé à une surface vibrante. Pour un corps donné, le rayonnement varie selon le mode vibratoire. En conséquence, le rayonnement se modifie continuellement selon les fréquences jouées.

Le son entendu à une certaine distance d'un instrument dépend du mode de propagation des ondes. En champ libre, par exemple en plein air, la propagation des ondes est sphérique (voir Glossaire, page 528). L'amplitude du front d'onde décroît en raison inverse du carré de la distance. S'y ajoutent les pertes dues à la transmission et l'absorption due à l'air, variable avec l'humidité. Dans un lieu clos, les ondes aériennes se réfléchissent sur les parois (murs, plafond, sol) et, si l'absorption est faible, il se forme rapidement un champ d'interférences plus ou moins homogène. Rappelons que la longueur d'onde λ est la distance entre deux points successifs dont les mouvements sont en phase. Dans un milieu donné, la longueur d'onde est inversement proportionnelle à la fréquence : elle diminue lorsque la fréquence croît. Par exemple, dans l'air à 15 °C, $\lambda = 3{,}40$ m pour un son de 100 Hz et $\lambda = 34$ cm pour un son de 1000 Hz. Comme le comportement des ondes (réflexion, diffraction) dépend des rapports entre λ et les dimensions des obstacles, la connaissance de ces ordres de grandeur permet d'estimer l'incidence que peuvent avoir certains obstacles sur le trajet des ondes sonores (voir Glossaire, page 523 et page 528). La tête d'un auditeur assis devant moi, d'un diamètre de l'ordre de 20 cm, n'est pas un obstacle pour les sons de fréquence basse, mais elle commence à faire écran pour les fréquences supérieures à 1700 Hz.

Le « son » d'un instrument, tel que nous nous le représentons en mémoire, est un concept abstrait. Dès qu'on veut faire une analyse, il faut garder à l'esprit le fait que le signal enregistré par un microphone est représentatif d'une position particulière de l'espace sonore. Il change d'un point à un autre.

2.6. La réception, la saisie du son, les transformations de la vibration

Ondes aériennes. Dans la situation ordinaire, nos oreilles captent les variations de la pression aérienne rayonnées directement par les vibrations des objets excités.

Capteurs (autres que l'oreille). Tout mouvement vibratoire peut être converti en signal électrique : il suffit de disposer du capteur approprié. Avec un accéléromètre, un capteur électromagnétique, un électroglottographe (voir chapitre 9, § 1.2.3) on peut capter directement le mouvement d'une structure vibrante, puis l'amplifier et le convertir en ondes aériennes. La guitare électrique, par exemple, est un instrument dans lequel le son rayonné directement par la structure (le son « acoustique ») ne joue qu'un faible rôle dans le résultat final. Le « son » que rayonne la guitare électrique s'élabore tout au long d'une chaîne dans laquelle interviennent des filtres, des modulateurs, des mises en forme dynamiques et, aujourd'hui, des interactions en temps réel avec des traitements informatiques. Même si l'origine en est le mouvement d'une corde pincée – capté par un système électomagnétique – et que parfois les vibrations solidiennes du corps de l'instrument et du manche y participent, le signal sonore résultant peut n'avoir aucune des caractéristiques acoustiques de la famille guitare. Quand au rayonnement dans l'espace, il est le résultat du mixage et de l'affectation d'une ou de plusieurs voies à un système d'enceintes. « L'écoute en est déportée et focalisée sur un autre objet que l'instrument lui-même »[12]. Le plus souvent maintenant, nous écoutons des sons préalablement enregistrés, reproduits par voie électroacoustique. Le signal sonore

12. Lähdeoja, O., Navarret, B., Quintans, S., & Sèdes, A., 2013, « La guitare électrique comme instrument augmenté et outil de création musicale », in *La musique et ses instruments*, Paris : Delatour, p. 317).

ainsi diffusé dans l'espace est assez différent de celui qu'une source mécanique rayonne par voie aérienne.

2.7. Définitions utiles

Soit un mouvement vibratoire constitué d'un aller-retour simple et régulier dans le temps : ce mouvement périodique est dit « sinusoïdal ».

La *fréquence* vibratoire (*f*) est le nombre d'oscillations qu'effectue le système (corde, anche) pendant une seconde. L'unité de fréquence est le hertz (Hz). La durée d'une seule oscillation porte un nom spécifique : c'est la *période* (T) du mouvement. En acoustique, les périodes, mesurées en secondes, sont souvent affichées en millisecondes (ms) par commodité.

La *période* (T) est l'inverse de la fréquence (*f*), soit T (s) = 1/*f* (Hz).

L'*amplitude* du mouvement vibratoire correspond au déplacement maximal de la portion excitée par rapport à sa position de repos. L'amplitude est liée à l'énergie injectée au moment de l'excitation.

L'*onde* décrit la propagation du mouvement vibratoire depuis le point d'impact jusqu'à l'ensemble de la structure. Elle possède une *célérité* c (m.s $^{-1}$) (vitesse de propagation) et une *longueur d'onde* λ (m) (périodicité spatiale) qui dépend de la fréquence et des caractéristiques du milieu de propagation selon la relation : λ (m) = c (m.s $^{-1}$)/*f* (Hz)

Voir Glossaire, pages 516 et 523.

Les oscillations sont *libres* (excitation par impulsion) ou *entretenues, périodiques* (excitation par entretien). Voir Glossaire page 526.

3. La forme temporelle des vibrations

3.1. L'inscription des vibrations

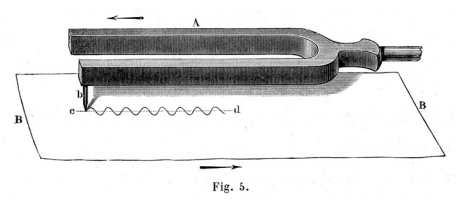


Figure 1.12 Inscription de la forme vibratoire d'une branche de diapason.

Helmholtz, H., 1874, *Théorie physiologique de la musique*, p. 27.

Dès les premières observations, les liens entre l'amplitude des vibrations et l'intensité sonore, entre la rapidité des vibrations et la hauteur des sons ont été établis. En revanche, l'étude de la forme des vibrations en un point donné, c'est-à-dire des relations entre la forme des vibrations et la qualité du timbre n'a préoccupé les scientifiques que plus tardivement, d'autant que l'observation des mouvements à l'œil nu était difficile en raison de la rapidité des vibrations et de leur faible amplitude. Il était commun d'imputer au matériau les différences de timbre[13]. « L'écriture » du mouvement de la branche d'un diapason (figure 1.12) est une opération difficile à réaliser. Si l'on veut, par exemple, étudier la période d'un diapason de 100 Hz (T = 10 ms), il faut, pour obtenir 1 cm de tracé, que le papier défile à 1 m/s avec une très grande régularité. Les nombreuses tentatives réalisées au milieu du XIXe siècle aboutiront à la

13. « Le différent timbre du son (sic) et ses articulations sont au nombre des objets les plus remarquables de l'ouïe. Elles ne paraissent pas dépendre des manières de vibrations, ni (ou très peu) de la forme du corps sonore, mais plutôt (§ 31) de la matière du corps sonore et de celle du corps par lequel il est frotté ou frappé, comme aussi de la matière qui propage le son. » Chladni, 1809, § 240.

conception d'appareils à « écrire le son », comme le phonautographe de Scott de Martinville (1857), et finalement à la gravure du son dans de la cire (Edison, 1877)[14].

3.2. La composition des vibrations

À l'époque où Helmholtz entreprend ses recherches, les inscripteurs à stylet sont très imparfaits[15]. Sur la base de la loi mathématique de Fourier (voir Glossaire, page 519), il démontre objectivement la présence d'harmoniques dans un son complexe périodique à l'aide de résonateurs. Helmholtz établit ainsi la première théorie acoustique de ce qu'il appelle le « timbre musical » qui ne concerne que la partie stable des sons instrumentaux entretenus. Il démontre aussi que la phase (voir Glossaire) relative des harmoniques n'a pas d'incidence sur la perception du timbre.

3.3. Du mouvement à l'onde sonore

Avec les progrès techniques (amplification électrique et oscillographe cathodique, électronique, numérisation), la captation et l'inscription des mouvements vibratoires permettent de reproduire avec fidélité les moindres détails de l'onde sonore.

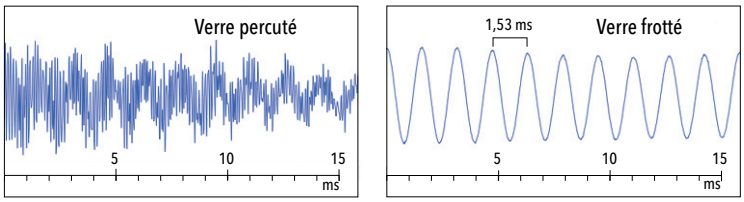


Figure 1.13 Variation d'amplitude en fonction du temps pour deux types d'excitation d'un verre. À gauche, *percussion*. Tous les modes propres sont excités simultanément. Le signal temporel est complexe. À droite, *frottement*. Le verre vibre à la fréquence du mode entretenu. Le signal temporel est régulier, périodique et l'on peut mesurer la période du mouvement indiquée sur la figure (ici 1,53 ms). Les signaux représentés sont captés 100 ms après le début de l'excitation.

On voit sur le tracé de l'amplitude de la figure 1.13 l'opposition entre l'onde périodique du signal émis par un verre frotté (à droite) et l'onde complexe, irrégulière du signal émis par le verre percuté (à gauche). Cette dernière est animée de petites oscillations correspondant aux fréquences des différents modes propres, non synchronisés. À l'aide du filtrage, nous proposons d'écouter les différentes composantes qui constituent chacun des deux sons émis par le verre (Sons **1.7** et **1.8**).

14. Voir Paul Charbon, 1981.
15. Helmholtz observe à l'œil nu les mouvements des corps vibrants (corde, diapason) à l'aide de points lumineux et d'un « microscope à vibration » (p. 113 de l'édition française, 1874).

4. L'analyse auditive des composantes d'un son : la série harmonique

4.1. L'analyse auditive par filtrage

Plusieurs techniques permettent d'isoler les composantes d'un son complexe. Celle qui est utilisée dans les exemples qui suivent a été réalisée avec un matériel analogique. Le son enregistré sur une boucle magnétique se répète indéfiniment. Il passe au travers d'un filtre à bande étroite dont la fréquence centrale est variable, ce qui permet de sélectionner les composantes les unes après les autres. La composante sélectionnée est amplifiée sans pour autant que les autres soient totalement éliminées. Cet effet, dû aux limites de performance des filtres analogiques offre un avantage auditif certain : le son global reste perceptible. Les techniques d'analyse-synthèse numériques permettent aujourd'hui d'obtenir un filtrage total.

4.2. Le verre percuté et le verre frotté

La « dissection » sonore que permet le filtrage nous fait pénétrer au cœur des agrégats les plus complexes. Dans le son du verre percuté (Son **1.7**), les fréquences émises individuellement par chacun des modes propres, qui ont des évolutions temporelles indépendantes, sont déjà perceptibles dans le son global. En particulier, nous pouvons prêter attention aux différentes *notes* qui émergent alors que d'autres s'éteignent. Après l'écoute des composantes isolées, le son global se laisse plus facilement analyser (voir figure 1.14).

Son 1.7 (31")

Filtrage : verre percuté

Il en va tout autrement de l'écoute du son entretenu (Son **1.8**). Le son global se présente comme un tout relativement simple à percevoir : *une seule note*, mais pourvue d'une sonorité particulière. La première composante isolée ne diffère du son global que par sa sonorité douce et mystérieuse : c'est un *son pur*. La suite du filtrage révèle une succession de sons de hauteurs diverses – les harmoniques – pratiquement inaudibles dans le son global, à l'exception de l'harmonique 5 qui se distingue par son intensité.

Son 1.8 (31")

Filtrage : verre frotté

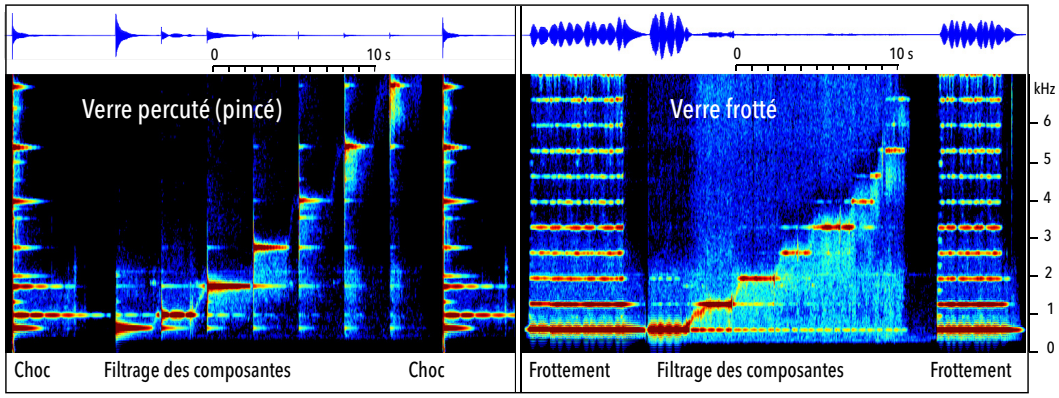


Figure 1.14 Cette figure illustre l'écoute des analyses par filtrage du Son 1.7 (verre excité par percussion) et du Son 1.8 (verre excité par frottement). Pour chaque exemple, on entend et on voit le son global, puis l'analyse par mise en résonance des composantes sélectionnées successivement du grave à l'aigu, et le son global qui réapparaît à la fin de la séquence. La courbe supérieure rend compte de la variation d'amplitude en fonction du temps.

Voici les cinq premiers sons de chaque suite (notes musicales approchées) :

Verre percuté	(1) mi_4	(2) si_4	(3) la_5	(4) fa_6	(5) si_6
Verre frotté	(1) mi_4	(2) mi_5	(3) si_5	(4) mi_6	(5) $sol\#_6$

Seul le premier son est commun aux deux séries : c'est la fréquence du mode 1.

Chaque composante émise par le verre percuté correspond à la fréquence d'un des modes propres de vibration du verre : cette suite est spécifique d'un verre donné. Au contraire, la suite des composantes du verre frotté est celle que l'on va retrouver pour tous les sons entretenus : c'est la *série harmonique*.

4.3. La série harmonique

> *La chorde frappée, & sonnée à vuide fait du moins cinq sons différens en mesme temps, dont le premier est le son naturel de la chorde, qui sert de fondement aux autres … Or il faut choisir un grand silence pour les appercevoir, encore qu'il ne soit plus necessaire quand on y a l'oreille accoustumée … & j'ay rencontré plusieurs Musiciens qui les entendent aussi bien que moy … Or ces sons suivent la raison de ces nombres 1, 2, 3, 4, 5, car l'on entend quatre sons differens du naturel, dont le premier est à l'Octave en haut, le second à la Douzième, le 3 à la Quinzième, & le 4 à la Dix-Septiesme majeure …*
>
> Marin Mersenne, 1636, Livre quatrième des instruments, Proposition IX.

Quelques auteurs comme Mersenne et Descartes[16] avaient déjà remarqué, dès le XVIIe siècle, que dans certaines conditions de silence, et particulièrement pour les sons graves (tuyau d'orgue, viole de gambe), il était possible d'entendre, en plus de la note fondamentale, des sons à la quinte redoublée (douzième) ou à la tierce majeure (dix-septième). Ce phénomène intrigant est resté sans explication jusqu'au XIXe siècle. C'est Joseph Fourier qui a montré (*Théorie analytique de la chaleur*, 1822) que l'on pouvait décomposer un mouvement périodique complexe en une somme de mouvements périodiques élémentaires dont les fréquences sont des multiples exacts de la composante la plus grave, appelée fondamentale. Ainsi les fréquences des cinq premiers harmoniques (H) d'un la_3 de 440 Hz seront : H1 = 440 Hz ; H2 = 2 × 440 = 880 Hz ; H3 = 3 × 440 = 1 320 Hz ; H4 = 4 × 440 = 1 760 Hz ; H5 = 5 × 440 = 2 200 Hz, et ainsi de suite. La succession des harmoniques d'un son périodique est invariable, c'est la même pour tous les sons périodiques.

Son 1.9 (1'45)

Harmonica
(voir page 31)

Son 1.10 (1'27)

Trombone

Son 1.11 (1'45)

Basson

Son 1.12 (17")

Voix diphonique
(voir page 31)

Voici quelques exemples d'analyse de sons instrumentaux : Son **1.9** (harmonica, $ré_2$) ; Son **1.10** (trombone, sol_1) et Son **1.11** (basson, sib_1). Plusieurs techniques vocales, telles le chant « diphonique » utilisent la sélection des harmoniques pour produire une mélodie. Écoutez l'analyse du Son **1.12**. Pour plus d'explications, voir page 26.

4.4. Les intervalles de la série harmonique

Par définition, l'intervalle entre deux sons A et B est déterminé par le rapport entre leurs fréquences, soit f(B)/f(A) (ou inversement). Étant donné que les harmoniques d'un son de fréquence f ont pour fréquences 2f, 3f, 4f etc., les intervalles entre deux sons successifs ont des rapports de fréquence qui sont : 2, 3/2, 4/3, 5/4, et ainsi de suite. Du point de vue musical, la mélodie formée par la suite des harmoniques est invariable et fournit toujours la succession des intervalles suivants : *octave, quinte, quarte, tierce majeure, tierce mineure, petite tierce mineure, ton majeur,* etc.

16. *Compendium musicae ou Abrégé de musique* (manuscrit 1618). Voir aussi Baskévitch, 2008.

Lorsqu'on connaît cette suite, il est facile d'associer aux intervalles les rapports numériques qui les caractérisent.

Prenons l'exemple d'un son dont la hauteur fondamentale est mi_1 (figure 1.15).

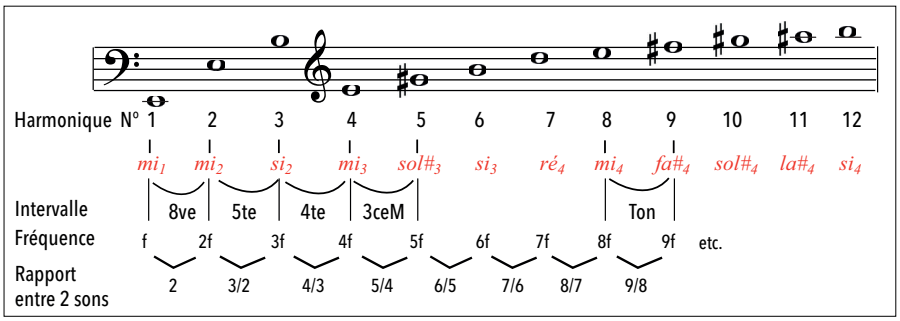


Figure 1.15 Harmoniques du mi_1. Notation musicale, nom des notes et indice d'octave. Intervalles et rapports numériques entre deux sons successifs.

La suite des huit premiers harmoniques correspond aux notes mi_1, mi_2, si_2, mi_3, $sol\#_3$, si_3, $ré_4$, mi_4.

- Remarque 1 : l'octave est un intervalle particulier de rapport 2. Les harmoniques n° 2, 4, 8, 16 sont donc tous à intervalles d'octave du premier. Pour la même raison, les sons harmoniques de numéros pairs sont toujours la réplique à l'octave supérieure d'un harmonique déjà apparu : par exemple, H3 et H6 ou H5 et H10. Finalement, seuls les harmoniques de rang impair introduisent un son nouveau dans la série.
- Remarque 2 : la notation musicale des harmoniques est très pratique. Il est utile de la connaître par cœur, au moins jusqu'à l'harmonique 12. Mais il faut garder à l'esprit que les notes écrites sur une portée (voir figure 1.15) sont des approximations. Il est courant d'entendre dire, à l'audition d'une suite d'harmoniques : « la tierce majeure (H 5) est trop basse ». L'harmonique 5 n'est ni juste ni faux, il est très exactement à la fréquence quintuple du fondamental. Il forme avec l'harmonique 4 un intervalle de tierce majeure *pure* ou *naturelle*, dont le rapport exact est 5/4. Cette tierce est plus faible que celle du tempérament égal qui est notre référence implicite d'évaluation de la justesse[17]. Les écarts entre les intervalles de la série harmonique et ceux des échelles musicales sont l'objet de discussions permanentes (voir chapitre 8, page 401).
- Remarque 3 : au fur et à mesure que l'on monte dans l'aigu la grandeur des intervalles entre les sons successifs de la série harmonique décroît régulièrement. À partir du 26e harmonique tous les intervalles sont inférieurs au demi-ton chromatique.

Inversement à l'analyse, on peut faire la synthèse harmonique d'un son complexe périodique par addition de sinusoïdes ayant des fréquences multiples les unes des autres (Son **1.13a** et **1.13b**).

Son 1.13a (28″)

Synthèse additive progressive

Son 1.13b (25″)

Écoute de l'harmonique isolé avant addition *(voir page 31)*

17. En toute rigueur, aucun des intervalles d'un piano, même bien accordé, ne correspond exactement à un intervalle de la série harmonique. En revanche, la quinte des cordes à vide d'un violoniste ou l'octave réalisée sur un orgue sont des intervalles purs, sans battement, comme ceux des harmoniques.

5. Vocabulaire : de la musique à l'acoustique et inversement

Avant de poursuivre, il importe de définir le sens de quelques termes utilisés en acoustique. Certains comme « fondamental » ont des sens multiples que l'on doit préciser à chaque fois. D'autres comme « harmonique » désignent en musique des phénomènes tout à fait différents. Le lecteur pourra aussi se reporter au Glossaire.

5.1. Les divers types de sons

Un *son pur* est produit par une vibration simple, sinusoïdale ; il ne comporte qu'une seule fréquence. Il est très facile de synthétiser un son pur, mais peu de sources naturelles en produisent, à l'exception du sifflet oral.

Tout son qui n'est pas pur est *complexe*. Chacune des composantes d'un son complexe périodique est un son pur. Un son complexe *périodique* est harmonique. Un son complexe *apériodique* est *inharmonique* ou *quasi périodique* lorsque les composantes ont des fréquences voisines de celles des harmoniques. C'est le cas du piano[18]. Les instruments à cordes frappées ou pincées ne sont pas (en toute rigueur) harmoniques. On définit un degré *d'inharmonicité* qui varie avec les caractéristiques des cordes (raideur). Nous sommes particulièrement sensibles à l'inharmonicité des sons graves de la harpe et du piano.

5.2. Les termes à connaître : harmoniques, partiels, fondamental

Ces termes ont des sens différents selon que l'on parle des composantes simultanées d'un son, ou que l'on considère les sons produits par les différents modes vibratoires d'un corps.

5.2.1. Les composantes simultanées d'un son isolé

Harmonique est un terme que l'on doit réserver pour désigner une composante d'un son périodique. Par définition, la fréquence d'un harmonique est un multiple entier de la fréquence de l'harmonique 1 ou *fréquence fondamentale*[19].

Son 1.14 (1'54)

Cloche d'église : analyse par filtrage

Partiel est un terme général qui peut désigner toute composante fréquentielle isolable du spectre d'un son quelconque, mais le plus souvent on l'utilise pour les sons non entretenus : les partiels d'une cloche, les partiels d'un son de piano. Le partiel le plus grave d'un son inharmonique est aussi appelé *fondamental*. Les fondeurs de cloche accordent les modes propres vibratoires de sorte que les partiels forment un accord mineur agréable à entendre. Écoutez l'analyse des partiels d'une cloche dans le Son **1.14** (voir aussi chapitre 6, page 255).

Le terme *fondamental* a donc plusieurs sens. Il arrive aussi que l'on désigne par fondamental le son produit par le premier mode propre d'un corps : on parle du *fondamental d'un tuyau* (qui est aussi le partiel 1 du tuyau).

Du point de vue perceptif, les partiels peuvent être perçus individuellement, alors que les harmoniques fusionnent en produisant la *hauteur fondamentale* du son, qu'il y ait ou non de l'énergie à la fréquence fondamentale correspondante (voir chapitre 6, page 238)

18. Dans la plupart des cas, les composantes des sons quasi périodiques ont des fréquences supérieures à celles des harmoniques des numéros correspondants. Elles sont « plus hautes » que les harmoniques et l'écart croît avec leur rang.
19. Il faut noter que, dans la communauté des chercheurs qui travaillent sur la parole, cette fréquence est désignée par f_0.

L'adoption d'un vocabulaire rigoureux en acoustique permet de discriminer clairement les deux catégories de production sonore que nous avons présentées. Un corps vibrant en oscillations libres émet un agrégat sonore dont chaque partiel provient de la fréquence d'un mode propre. Un corps dont la vibration est entretenue périodiquement sur un mode propre donné émet un son composé d'harmoniques.

5.2.2. Les sons successifs produits par les modes vibratoires d'un corps

Partiels d'un tuyau ou d'une corde

Lorsque le corps vibrant est long et fin (corde, tuyau), les intervalles entre les partiels des modes vibratoires successifs ont beaucoup de ressemblance avec ceux d'une série harmonique[20], d'où l'emploi du terme *harmonique* (au lieu de partiel) par certains instrumentistes, ce qui produit une certaine confusion.

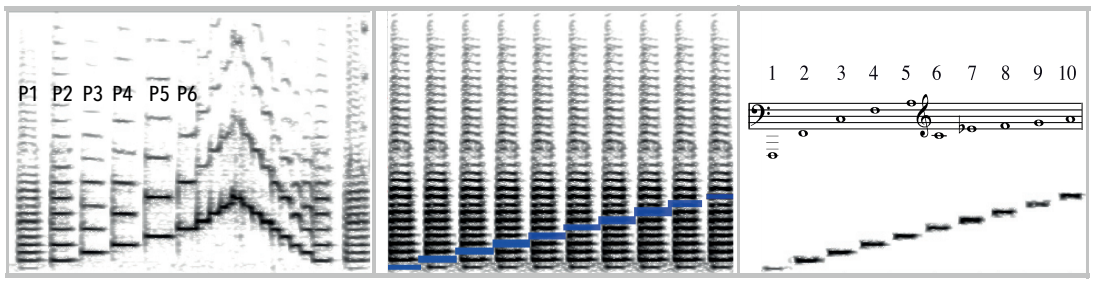


Figure 1.16 Tuba, doigté fa_0. (A) - Analyse du jeu des dix premiers partiels du tube, en série ascendante puis descendante. Chaque partiel est un son riche en harmoniques. (B) - Le premier partiel du tube a été copié dix fois pour réaliser le filtrage des harmoniques. (C) - Extraction des dix premiers harmoniques du premier partiel du tube (logiciel Audiosculpt).

Prenons pour exemple le tuba, qui permet de jouer la série complète des modes vibratoires du tube (Son **1.15**). Sur la position fa_0, l'instrumentiste joue successivement les notes fa_0, fa_1, do_2, fa_2, la_2, etc., sans changer de doigté, chaque note correspondant à un mode vibratoire particulier du tuyau. Sur la figure 1.16 (A), on voit bien que chaque partiel du tuyau est un son complexe ayant ses harmoniques. L'emploi du mot partiel correspond ici au fait que, pour chaque mode, la configuration vibratoire du tuyau présente plusieurs *parties*.

Son 1.15 (8")

Tuba : jeu des partiels 1 à 10 (doigté fa_0)

Pour bien marquer la différence entre *partiel du tuyau* et *harmonique d'un son,* nous avons effectué le filtrage des harmoniques du premier partiel fa_0 (C). À l'écoute du Son **1.16**, on reconnaît la sonorité caractéristique des sons purs d'une série harmonique distincte de celle des partiels successifs du tuyau.

Son 1.16 (7")

Tuba : filtrage des harmoniques 1 à 10 du partiel 1 (note fa_0)

Dans le cas du trombone ténor (Son **1.22**), la note sib_2 est déjà le deuxième partiel du tube, car l'intervalle entendu entre les deux premiers sons est une quinte et non une octave.

La flûte *octavie* parce que le partiel 2 est pratiquement à l'octave supérieure du fondamental, ou 1er partiel (son **1.17**). La clarinette n'octavie pas, elle *quintoie*, car le partiel 2 est à la douzième[21] (octave + quinte) du partiel 1 et non à l'octave, ce qui est aussi le cas des tuyaux de flûte bouchés (bourdon d'orgue, flûte de pan)[22]. Il existe d'ailleurs une relation directe entre les fréquences des modes propres et le contenu spectral des sons qui fait que nous reconnaissons à l'oreille la sonorité d'un bourdon ou celle des

Son 1.17 (11")

Partiels 1 et 2 : flûte traversière puis clarinette. *(voir page 31)*

20. L'ajustement des intervalles entre les modes vibratoires successifs est le fruit d'une longue expérience des facteurs.
21. Douzième degré de l'échelle diatonique partant du fondamental. « Douzième, Quinzième et Dix-Septième » sont des termes musicaux anciens.
22. Voir plus loin (Son 1.25) le jeu de la tilinca qui utilise les deux séries : tuyau ouvert et tuyau bouché.

notes graves de la clarinette (les partiels du premier mode) : les harmoniques impairs, dont les fréquences correspondent à celles des modes propres ont beaucoup plus d'intensité. C'est le contraire pour les instruments qui octavient.

Les *harmoniques effleurés* que l'on joue sur les instruments à cordes (violon, harpe) sont en réalité les sons *partiels des modes propres* de la corde. Le musicien excite la corde sur un mode vibratoire différent du mode habituel, qui est le premier mode. Un harmonique effleuré est un nouveau son **fondamental** qui possède des harmoniques s'il est entretenu (violon, flûte) ou ses propres partiels s'il est en oscillations libres (guitare, harpe, pizzicati de violon). Écoutez la suite des partiels d'une corde de guitare (Son **1.18**).

Son 1.18 (48")

Partiels 1 à 8 : corde de guitare

Harmoniques inférieurs (sous harmoniques)

5.2.3. La série harmonique inverse et le period-doubling

L'idée d'une série harmonique inverse, descendante, générant une suite de *sous-harmoniques* circule dans divers ouvrages de musique et sert de justification à l'existence du mode mineur. Ce concept n'a pas de réalité physique. En revanche, des recherches récentes ont permis de mettre en évidence des comportements vibratoires non linéaires donnant lieu à la production de fréquences fondamentales plus basses que le premier mode habituel. Il s'agit généralement de l'octave inférieure[23] et, plus rarement, de la quinte. La pratique en est recherchée dans certaines techniques vocales : chant tibétain, chanteuses xhosas d'Afrique du Sud, voix de Bassu des Sardes. Ces phénomènes (que l'on rencontre aussi dans certaines dysphonies) sont généralement évités dans l'esthétique classique. Ajoutons qu'un son en *period-doubling* (donc de fréquence moitié) possède, comme toute vibration périodique, une série harmonique normale, c'est-à-dire ascendante !

6. Les mélodies spectrales et les mélodies de partiels successifs : exemples musicaux

Un grand nombre de traditions musicales ont développé de par le monde l'usage de techniques de jeu basées soit sur la sélection des harmoniques d'un son par résonance, soit sur la série des partiels correspondant aux modes vibratoires successifs d'une structure.

6.1. Les mélodies spectrales : sélection d'harmoniques par résonance buccale

Plusieurs techniques musicales utilisent la sélection d'harmoniques par la cavité buccale pour produire une mélodie. Celle-ci est un résonateur assez efficace, ajustable en dimensions et d'usage universel. Sa zone d'action se situe entre 500 et 2000 Hz[24]. Dans le chant diphonique (ou harmonique) cette sélection s'opère directement sur le spectre harmonique de la voix du chanteur qui maintient généralement la fondamentale constante, comme dans le Son **1.12** (voir page 22). Une grande diversité de techniques est pratiquée dans le monde, et notamment en Mongolie. Écoutez le Son **1.19** (voir aussi le chapitre 6, § 2.7, page 244 et le chapitre 9, § 3, page 462).

Son 1.19 (30")

Chant diphonique

23. Phénomène connu sous le nom de *period doubling*, analysé au chapitre 9, page 464.
24. Le son que l'on produit en sifflant correspond à la fréquence de résonance de la cavité buccale, exactement comme lorsqu'on souffle sur le bord d'une bouteille. La note la plus grave (langue aplatie, très en arrière) ne descend guère au dessous de 500 Hz). Le son le plus aigu produit avec la même technique (langue très en avant, presque sur les dents) monte à 3000 Hz et plus.

Figure 1.17 Jeu de la guimbarde.
Cliché M. C.

Le jeu de la guimbarde résulte d'un couplage entre la cavité buccale et une lame vibrante excitée de façon impulsive (figure 1.17). En toute rigueur, il s'agit de partiels quasi harmoniques. Dans l'exemple sonore **1.20**, John Wright présente les interactions entre l'instrument et le joueur et alterne mélodie chantée ou jouée à la guimbarde.

Son 1.20a (1'05")
Démonstrations du jeu de la guimbarde

Son 1.20b (1'14")
Musique de guimbarde
Voir page 32

L'échelle mélodique utilisable avec ces techniques de jeu est strictement celle des intervalles entre les harmoniques (intervalles approchés pour les partiels). Nous avons vu que les intervalles musicaux entre les premiers harmoniques sont grands : octave, quinte, etc. Si l'on veut jouer un intervalle d'un ton, analogue à celui de notre système diatonique, on ne le trouve qu'entre les harmoniques 8 et 9. Plus on monte dans le rang des harmoniques et plus les intervalles se resserrent : il devient alors difficile de sélectionner à coup sûr un son précis, étant donné que la zone d'action en fréquence de la cavité buccale est réduite. Le fondamental vocal ou instrumental doit être choisi en fonction du plus petit intervalle souhaité.

Deux fondamentaux successifs

Certains musiciens résolvent le problème de façon élégante en jouant avec deux fondamentaux à intervalle d'un ton. Prenons comme exemple les harmoniques de deux sons do_2 et $ré_2$.

Tableau 2. Harmoniques de deux sons à intervalle d'un ton

H1	H2	H3	H4	H5	H6
do_2	do_3	sol_3	do_4	mi_4	sol_4
$ré_2$	$ré_3$	la_3	$ré_4$	$fa\#_4$	la_4

On voit que, dès le troisième harmonique on peut jouer, en changeant de fondamental : sol_3, la_3, do_4, $ré_4$, mi_4, $fa\#_4$, sol_4, avec les harmoniques 3, 4, 5, 6 de chacune des deux séries.
Écoutez *l'arc musical* ngbaka (figure 1.18), de l'exemple sonore **1.21** dont les deux notes fondamentales sont sol_2 (pour la corde à vide) et la_2 (pour la corde raccourcie). Voir aussi le chapitre 6, page 247 et le chapitre 8, page 402.

Figure 1.18 Jeu de l'arc musical. La corde passe entre les lèvres du musicien ; il en raccourcit la longueur à l'aide du bâtonnet de la main gauche.
Cliché M. C. Musicien, Michel Kossi, groupe Ndima.

Son 1.21 (31")
Arc musical

6.2. Les mélodies produites avec les partiels d'un tuyau ou d'une corde

6.2.1. Les instruments à vent

Son 1.22 (8″)

Trombone à coulisse

Le jeu des instruments à embouchure, ou cuivres, consiste à sélectionner la succession des sons partiels du tube. L'échelle musicale de ces instruments est directement liée à la justesse relative des modes, laquelle dépend principalement de la perce du tube (à l'exception du trombone à coulisse qui peut s'ajuster en modifiant la position de la coulisse). On apprécie la justesse en comparant les intervalles entre les partiels successifs avec ceux de la série harmonique du son le plus grave qui sert de référence. Celui-ci, très faible, est rarement joué et le jeu mélodique exploite le plus souvent les partiels 2 à 12, comme dans l'exemple du Son **1.22** joué au trombone à coulisse ténor (sans barillet).

Son 1.23 (31″)

Cor des Alpes

Son 1.24 (20″)

Trompe de chasse

Citons, parmi ces instruments, les cors (dont l'ancien cor d'orchestre dit « à tons de rechange »), les trompes, la trompette de cavalerie, le clairon. Écoutez aussi les exemples du cor des Alpes (Son **1.23**) et de la trompe de chasse (Son **1.24**). La trompe de chasse est reconnaissable par l'usage du 11e partiel qui forme un intervalle de quarte augmentée avec la tonique *fa*. Dans le jeu des instruments de tradition occidentale le partiel 7 n'est pas utilisé : les musiciens *sautent* du son 6 au son 8. Marin Mersenne tente, dans un paragraphe entier, de trouver des explications à cette « anomalie »[25]. Allez au chapitre 8, § 1.3, page 398 pour une discussion sur l'harmonique 7.

Son 1.25 (50″)

Flûte roumaine tilinca

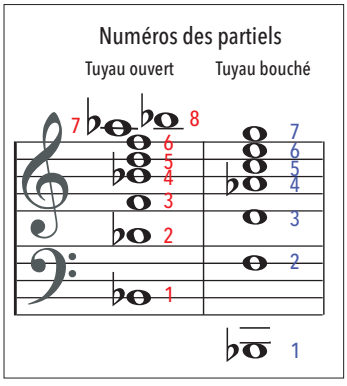


Figure 1.19 Séries de partiels de la flûte tilinca : tuyau ouvert et tuyau bouché.

Certaines flûtes longues et fines, sans trous latéraux comme la *tilinca* roumaine, produisent aussi des mélodies sur la suite des partiels (Son **1.25**). Dans cet exemple, le musicien joue sur deux séries car, en obturant l'extrémité inférieure du tuyau avec un doigt, il obtient une autre série dont les sons, qui correspondent aux modes propres d'un tuyau bouché, s'intercalent entre ceux de la première série (voir l'analyse spectrale au chapitre 2, § 4.6, page 65). Cette technique de jeu a plusieurs avantages. La combinaison des deux séries fournit une échelle complète de seize sons alors que le flûtiste ne travaille que sur les partiels 3 à 8 de chaque série (au delà il est de plus en plus difficile de sélectionner un partiel au coup de langue). Le bruit de bouchage du tuyau joue un rôle rythmique.

Les deux séries intercalées sont présentées sur la figure 1.19 et ci-dessous : tuyau bouché (en italique), tuyau ouvert (en gras).

sib$_0$ **sib**$_1$ *fa*$_2$ **sib**$_2$ *ré*$_3$ **fa**$_3$ *lab*$_3$ **sib**$_3$ *do*$_4$ **ré**$_4$ *mi*$_4$ **fa**$_4$ *sol*$_4$ **lab**$_4$ *la*$_4$ **sib**$_4$

25. Voir l'*Harmonie universelle,* 1636, Livre troisième des instruments, Proposition XIV (page 252 Édition CNRS). *« Expliquer pourquoy la Trompette ne fait pas la Sesquisexte dans son cinquiesme intervalle, et qu'elle quitte le progrez qu'elle avoit suivy iusque au sixiesme ton pour faire la Quarte qu'elle avoit desia faite au troisiesme intervalle ».* Une des explications de Mersenne ne manque pas de poésie : *« L'on peut encore dire que la nature ayant donné les six tons, comme ses six iournées ausquelles elle se repose, qu'elle imite son Autheur qui se reposa à la fin des six iours ».*

6.2.2. La trompette marine

Parmi les instruments à cordes, citons la trompette marine qu'affectionne le *Bourgeois gentilhomme* de Molière.

L'instrument ne comporte qu'une corde, effleurée en divers endroits par la main gauche (voir figures 1.20 et 1.21). Observez la position de l'archet qui est situé entre le sillet et la main gauche. La mélodie utilise la suite des partiels de la corde (ou harmoniques effleurés) dont le son est considérablement transformé, voire distordu, par un chevalet instable, analogue à celui du « chien » de la vielle à roue (Son **1.26**) (voir Engel, 1992 ; Leipp, 1965).

Son 1.26 (30″)

Trompette marine

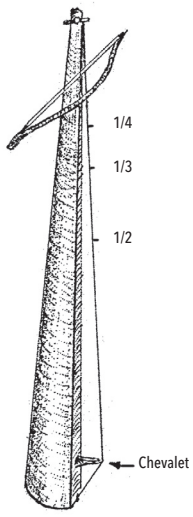


Figure 1.20 La trompette marine est un monocorde à archet dont l'échelle mélodique, obtenue par division de la corde, est celle de la suite des partiels. Comme pour le cor naturel et les flûtes de type tilinca, les intervalles sont voisins de ceux de la série harmonique.

Leipp, E., 1965, *Bulletin du GAM* n° 12.

Figure 1.21 Un joueur de trompette marine.

Colomb, C., 1878, *La Musique*, figure 113.

6.2.3. Le monocorde vietnamien

De façon similaire à la trompette marine le joueur de monocorde vietnamien (figure 1.22) joue les « harmoniques effleurés » de la corde. Une explication détaillée est donnée par le musicien dans l'exemple sonore **1.27**.

Son 1.27 (33″)

Jeu du monocorde, explications par Trân Van Khe

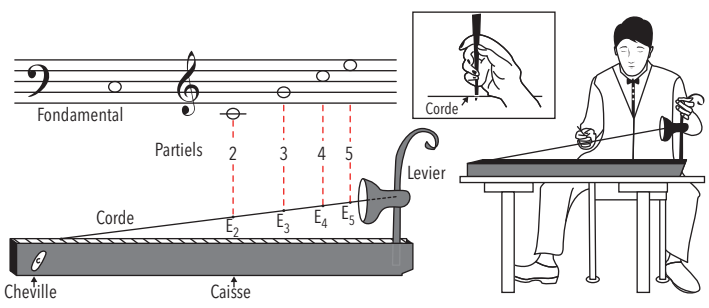


Figure 1.22 Le monocorde vietnamien. Le musicien pince la corde métallique au moyen d'un stylet de bois et, dans le même temps, il immobilise la corde à un nœud de vibration. Il obtient donc la suite des harmoniques effleurés de la corde. En changeant la tension par le jeu de la main gauche, il varie la fréquence fondamentale dans un intervalle d'octave. Pendant l'extinction du son, le musicien produit des ornements par de rapides mouvements de la main gauche.

Trân Van Khe., 1965, *Bulletin du GAM* n° 12.

Son 1.28 (22″)

Monocorde : musique

Pour produire les modes propres le musicien effectue de la main droite un geste qui enchaîne rapidement trois actions : l'effleurement de la corde à un point bien précis (production d'un nœud vibratoire), le pincement et la levée de la main. De la main gauche, il varie la tension de la corde en agissant sur la tige souple à laquelle elle est fixée. Cet ingénieux système confère à l'instrument des possibilités mélodiques et ornementales infinies. La musique produite n'est plus du tout dépendante des sons de la série harmonique puisque la hauteur peut être modulée par les changements de tension produits par la main gauche et le répertoire des ornementations que l'on joue pendant l'extinction du son est d'une grande richesse (Son **1.28** et figure 1.23).

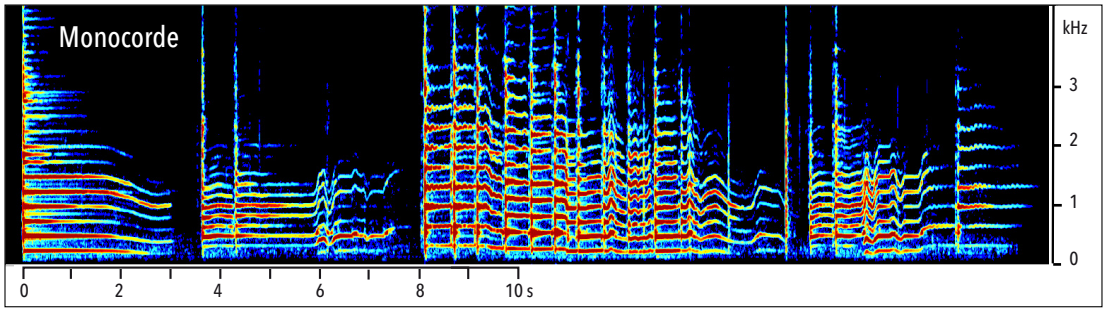


Figure 1.23 Sonagramme du jeu du monocorde (Son 1.28). Chaque nouvelle attaque de la corde correspond à un trait vertical suivi de raies harmoniques. Les dessins mélodiques sont réalisés par les changements de tension de la corde pendant l'extinction du son.

7. Les sons du chapitre 1

Remarque : certains exemples, le son 1.3 et les sons filtrés dont la première composante est grave (trombone, basson, tuba, cloche), ne sont audibles qu'avec une bonne qualité d'écoute.

7.1. Types d'excitation et modes vibratoires

Son 1.1 – Son du verre percuté. [M. C.]

Son 1.2 – Son du même verre frotté au doigt. [M. C.]

Son 1.3 – Vibraphone. Lame fa_2 percutée puis jouée à l'archet. [Base de données McGill]

Son 1.4 – Harpe. Corde mi_2, pincée au doigt puis entretenue avec un archet. [M. C.]

Son 1.5 – Cor des Alpes. Signal recueilli au pavillon d'un cor des Alpes qu'on excite par une impulsion à l'embouchure. Fréquence 46,72 Hz (environ fa_0). [M. C.]

Son 1.6 – Diapason à fourche. L'instrument est frappé puis posé sur une table. On entend d'abord un son très aigu qui est le deuxième mode vibratoire du système puis le premier mode vibratoire, le la_3 proprement dit, amplifié par couplage avec la table. [M. C.]

7.2. Analyses auditives des composantes d'un son

Son 1.7 – Verre percuté. Successivement : son global (mi_4) ; analyse par filtrage analogique des premiers partiels ; son global. [M. C.]

Son 1.8 – Verre frotté au doigt. Successivement : son global (mi_4) ; analyse par filtrage analogique des dix premiers harmoniques ; son global. [M. C.]

Son 1.9 – Harmonica. Le son $ré_2$ est mis en boucle (10 s.) puis filtré. On entend la succession des harmoniques amplifiés isolément dans l'ordre ascendant. À partir de l'harmonique 10 ($fa\#_5$), particulièrement intense, il devient difficile de les séparer. Filtrage descendant, puis le son global est donné à entendre de nouveau. [M. C.]

Son 1.10 – Trombone à coulisse, note sol_1. Séquence d'un son répété quatre fois de suite par le musicien, mise en boucle et filtrée. On entend successivement : le son original ; H1 (4 fois) ; H2 (4 fois) ; H3 (4 fois) ; H4 (4 fois) ; H5 à H16 répété 2 fois ; puis la suite balayée rapidement vers l'aigu ; pour finir le son global. [M. C.]

Son 1.11 – Basson, note sib_1 (117 Hz) avec vibrato. Successivement : le son répété deux fois ; puis filtrage ascendant des harmoniques H1 à H8 (une fois) ; montage rapide jusque vers 3 kHz; son global à la fin (deux fois). [M. C.]

Son 1.12 – Chant diphonique. Exemple didactique d'un glissando spectral ascendant et descendant produit sur un son vocal fixe. L'exemple normal (do_2, diapason haut) puis le même exemple ralenti à demi-vitesse et entendu à l'octave inférieure. La technique de chant diphonique a pour effet de produire un *filtrage par résonance*, tout à fait analogue aux exemples précédents produits avec un filtre analogique. Exemple chanté par Trân Quang Hai. [M. C.]

7.3. Exemples de synthèse additive numérique d'un son périodique

Son 1.13 – Deux exemples de synthèse d'un son composé de 10 harmoniques. Son 1.13a Synthèse n° 1. Successivement le son global, puis addition progressive des harmoniques (1, 1 + 2, 1 + 2 + 3, etc.). Son 1.13b Synthèse n° 2. Le son global, puis présentation isolée de l'harmonique avant son addition (1 ; 2 et 1 + 2 ; 3 et 1 + 2 + 3 ; 4 et 1 + 2 + 3 + 4, etc.). Son global à la fin. [M. C.]

7.4. Partiels et harmoniques

Son 1.14 – Cloche sonnée *à la volée* (4 coups). Filtrage des différents partiels, du grave à l'aigu (4000 Hz) puis de nouveau le son global à la fin (5 coups). Les partiels de cette cloche sont assez bien accordés mais ne suivent pas du tout la série harmonique. [M. C.]

Son 1.15 – Tuba : la suite des partiels obtenus sur le doigté du fa_0. Annonce, puis jeu des dix premiers partiels en série ascendante puis descendante. [G. Bucquet, LAM]

Son 1.16 – Tuba : les harmoniques du partiel 1, fa_0. Écoute des 10 premiers harmoniques du partiel obtenus par filtrage numérique (logiciel Audiosculpt). [M. C.]

Son 1.17 – Octaviation puis quintoiement. Flûte traversière en *ut*, doigté tout bouché : les deux premiers partiels du tuyau (do_3, do_4). Clarinette en *sib*, doigté tout bouché : les deux premiers partiels du tuyau ($ré_2$, la_3). Voir chapitre 2, figure 2.32 pour l'analyse. [Joe Wolfe, site Internet]

Son 1.18 – Harmoniques effleurés (ou partiels) d'une corde de guitare, *ré*$_2$. La corde à vide – qui est le partiel 1 – puis les partiels suivants, *ré*$_3$, *la*$_3$, *ré*$_4$, etc., E. Pélissier. [V. Mons]

7.5. Mélodies d'harmoniques : exemples musicaux

Son 1.19 – Chant diphonique (homme). Fondamental de la voix : *la*$_2$. *Chants de Mongolie*, n° 6, 1989 ; chanteur : T. Ganbold. [CD Auvidis W 260009]

Son 1.20 – Jeu de la guimbarde. Son 1.20a : 1re guimbarde (fondamental 59 Hz, *sib*$_0$ – 20 cents). Le musicien donne successivement les « notes » de la guimbarde seule, du grave à l'aigu, puis une petite *mélodie avec la bouche seule*, et de nouveau avec la guimbarde. Son 1.20b : 2e guimbarde (fondamental 71,8 Hz, *ré*$_1$ – 40 cents). Mélodie alternativement jouée à la guimbarde (certains passages sont sans souffle) et chantée. À noter : la voix chante trois octaves plus bas que la mélodie harmonique de la guimbarde. Exemples donnés par John Wright : disque 33 t., Spécial instrumental, *la guimbarde*, plage 1. [*Le Chant du Monde,* LDX 74434]

Son 1.21 – Arc musical. Jeu rapide et rythmé sur deux fondamentaux : *sol*$_2$ et *la*$_2$. Arc ngbaka, joué par N. Massemokobo (Centrafrique). [S. Arom, LAM, 1967]

7.6. Mélodies de partiels (tube ou corde)

Son 1.22 – La suite des partiels du trombone à coulisse, sur la position *sib*. Jeu des partiels dans l'ordre ascendant et descendant. Le premier son joué (*sib*$_1$), est le second partiel du tube. Musicien B. Sluchin. [Archives LAM]

Son 1.23 – Cor des Alpes : mélodie enregistrée en Suisse au cours d'un concours. Le fondamental est un *fa*. La mélodie commence sur le partiel 5 (tierce), ce qui se confirme par la succession des trois sons « *fa, sol, la* » dont les intervalles (tons) se situent entre les partiels 8, 9 et 10. On remarquera que le partiel 7 n'est pas utilisé. [Document W. Aebi, LAM]

Son 1.24 – Partiels de la trompe de chasse sur un *mib*. Le premier son joué, *mib*$_1$ (80 Hz), est déjà le deuxième partiel du tuyau. Le partiel 7 n'est pas joué par le sonneur (M. Pietri). On entend donc successivement les partiels 2, 3, 4, 5, 6, 8, 9, 10, 11, 12. [LAM]

Son 1.25 – Flûte roumaine tilinca. Le musicien joue sur deux séries de modes propres : celle du tuyau ouvert (fondamental *sib*$_2$) et celle du tuyau bouché (fondamental *sib*$_1$) *Anthology of Rumanian Folk Music* ; coffret 33 t. disque n° 1, plage 3a. [Electrecord EPD 78]

7.7. Instruments à cordes fonctionnant sur la suite des partiels

Son 1.26 – Trompette marine. Fondamental de la corde : *do*$_2$ = 130 Hz. Le premier son est le partiel 4 (double octave). Quand le *ré* apparaît, c'est le partiel 9. Extrait de *Guide des instruments baroques* ; T. Kosteletztki (1660-1722), joué par Max Engel. [Ricercar 93000]

Son 1.27 – Monocorde vietnamien. Technique de jeu présentée par M. Trân Van Khe (1965). Le « premier son », annoncé aussi « premier nœud », correspond au partiel 2 de la corde. [Archives LAM]

Son 1.28 – Monocorde vietnamien. Court exemple musical joué par M. Trân Van Khe. Notez l'importante variation de la fréquence de jeu d'un partiel donné obtenue par l'action de la main gauche sur la tension de la corde (dans cet exemple, on entend la descente à la quinte inférieure et la montée à la quarte supérieure). [Archives LAM]

CHAPITRE 2

LA REPRÉSENTATION DES SONS

1. Les analyses acoustiques et l'écoute

1.1. L'inscription des ondes sonores

De même que les odeurs, les phénomènes sonores sont fugitifs et insaisissables : ils s'évanouissent rapidement, ne laissant dans la mémoire de celui qui écoute qu'une trace partielle qui dépend de ses motivations et de ses capacités auditives. Seule une analyse physique du signal sonore effectivement produit peut servir de référence objective, de support d'échange entre plusieurs auditeurs, musiciens et scientifiques. Mais comment saisir le mouvement des ondes aériennes porteuses du son ?

Pendant des siècles, les sons dits « musicaux », c'est-à-dire les sons périodiques de hauteur définie, ont été l'objet de l'attention des philosophes, des mathématiciens et des physiciens qui ont progressivement dégagé les notions de fréquence, d'amplitude et de phase des mouvements vibratoires en étudiant la production mécanique des sons, puis ont accédé à la décomposition des vibrations périodiques complexes en mouvements élémentaires sur la base du théorème de Fourier. Les premières expériences d'analyse harmonique de sons musicaux réalisées par Helmholtz ont été effectuées à l'oreille, en direct, à partir de dispositifs sonores stables : tuyau d'orgue, corde frottée. À cette époque les connaissances acoustiques relevaient principalement de la mécanique des systèmes vibrants et bien peu des ondes sonores, que l'on ne savait pas comment capter. On doit la première tentative à un ingénieux typographe, Scott de Martinville[1], qui avait été fasciné par les planches d'un traité d'anatomie de l'oreille et se lança dans la construction d'un appareil captant les sons de façon similaire. Dès 1852, il produisit les premiers tracés d'ondes sonores sur une plaque de verre enfumée. Avec son phonautographe, ou « écriveur de sons », Scott était animé par l'espoir d'offrir à ses contemporains un procédé d'inscription directe de la parole, sans passer par le codage d'un alphabet écrit (voir historique, page 74). Il dut malheureusement se rendre à l'évidence : le tracé des ondulations n'était pas lisible à première vue.

Le signal obtenu par Scott contenait pourtant l'essentiel de l'information sonore, puisqu'il suffisait de le convertir à nouveau en ondes sonores pour entendre les séquences les plus complexes de musique ou d'environnement sonore, comme allait le montrer Edison vingt-cinq ans plus tard. Ce qui faisait défaut à Scott pour lire les courbes produites par son « oreille », c'est justement l'analyse. Il a fallu plus d'un siècle pour parvenir à l'analyse des ondes sonores en levant les obstacles techniques de la captation – en particulier les problèmes d'inertie et de sensibilité des capteurs – et en construisant des analyseurs spécialisés : filtres, enregistreurs de

1. Édouard-Léon Scott de Martinville, 1817-1879, inventeur du phonautographe. Voir Charbon, 1981, *La machine parlante*.

niveau, spectrographes. Aujourd'hui, les techniques numériques rendent possible l'application de méthodes mathématiques de décomposition des ondes, méthodes développées au sein d'une discipline à part entière : le traitement du signal. Le travail du son est devenu d'une subtilité et d'une précision telles que la resynthèse d'un son est identique à l'original : il est maintenant possible de faire des expériences sur le son comme on le fait avec une substance chimique.

Pourtant, de l'analyse physique des ondes sonores à la perception humaine, il reste encore un long chemin à faire. La première manipule des paramètres (fréquence, amplitude, durée), alors que la seconde se préoccupe de l'identité des sons et de leurs qualités.

La question qui se pose est donc la suivante : qu'attend-on de l'analyse acoustique ? Différentes mesures des paramètres du son ou bien une représentation de l'allure des phénomènes interprétables par un auditeur ? Ces deux attitudes antagonistes imposent des choix en amont assez opposés. Compte tenu du fait que les sons qui nous entourent sont « vivants », donc évolutifs, il est tout aussi important d'estimer leurs tendances ou leurs limites de variation dans un intervalle de temps donné que d'en mesurer précisément les paramètres à un instant donné plutôt qu'à un autre. En effet, il faut bien « arrêter » le temps pour effectuer la mesure et, ce faisant, le risque est grand de négliger des éléments importants pour l'interprétation. Pour ce qui nous préoccupe dans cet ouvrage les mesures sont nécessaires et leur précision souhaitable, à condition qu'elles aient un sens pour l'auditeur.

Lorsque la finalité de l'analyse acoustique est la mesure, il faut pouvoir contrôler toutes les étapes de l'opération, à commencer par la prise de son. Il est nécessaire d'utiliser un matériel professionnel, de procéder à des étalonnages rigoureux, d'enregistrer dans un lieu neutre comme une chambre anéchoïque, etc. La métrologie des sons est une discipline à part entière, extrêmement exigeante et finalement assez différente de ce qui nous préoccupe ici.

Notre objectif vise plutôt les représentations interprétables par des auditeurs. Or, contrairement à ce que d'aucuns pensent, effectuer une analyse n'est pas une opération automatique et neutre, car sa réalisation implique d'effectuer des choix dans les dimensions à représenter et dans les paramètres du calcul pour tenir compte de ce qui est pertinent perceptivement. D'une façon générale, il y a lieu de privilégier les méthodes représentant les variations des paramètres en fonction du temps, au détriment d'autres fournissant des données peut-être plus précises mais faisant l'impasse sur la dimension temporelle.

1.2. La représentation des sons

Une des difficultés de l'analyse tient au fait que la représentation exhaustive des phénomènes sonores (figure 2.1) nécessite au moins trois paramètres : l'amplitude, la fréquence et le temps[2]. Même si l'ordinateur permet maintenant de manipuler des graphiques en trois dimensions et de les faire tourner dans l'espace, et même si quelques sons particuliers peuvent bénéficier d'une telle représentation, la pratique commune privilégie encore les représentations dans un plan, avec seulement deux dimensions. Il faut donc combiner deux à deux les dimensions et travailler avec trois plans différents : amplitude/temps, fréquence/temps, amplitude/fréquence.

2. Nous laisserons provisoirement de côté la phase des signaux qui joue principalement un rôle dans la perception spatiale.

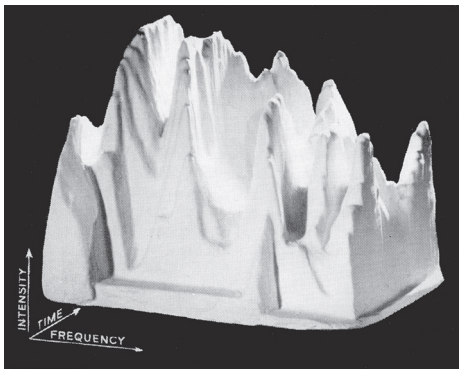


Figure 2.1 Une des plus anciennes représentations d'un son en 3D : le mot *nine* prononcé par une voix d'homme (Bell Laboratories).

Moles, A., 1960, planche II, page 33.

D'autres difficultés tiennent au fait que le son qui parvient aux oreilles du musicien portant un jugement n'est pas le même que celui qui est enregistré pour l'analyse, car le microphone qui capte le son n'est généralement pas placé au même endroit que son oreille. Cette différence affecte principalement l'amplitude des ondes sonores, donc la composition spectrale, l'analyse de la sonorité et, quelquefois aussi, la sensation de hauteur.

Finalement, avec un peu d'expérience, un musicien parvient à faire le lien entre les mesures de la fréquence, l'analyse du spectre et des transitoires et les qualités de hauteur et de sonorité perçues, car les sons instrumentaux sont déjà dimensionnés et catégorisés en vue d'une combinatoire musicale selon ces deux paramètres. Le problème se complique beaucoup lorsqu'on aborde la voix. Qu'écoute un auditeur d'opéra : les paroles ? la mélodie et son vibrato ? les « qualités » de la voix du chanteur ? l'émotion que celui-ci communique ? Certes, tout cela est dans le son, mais, au moment de passer à l'analyse, l'auditeur ne dispose que de l'amplitude, de la fréquence et de leurs évolutions dans le temps pour en rendre compte.

Malgré ces difficultés, nous faisons largement usage des analyses acoustiques dans les chapitres qui suivent, car la représentation des paramètres du signal sonore est un outil précieux par son objectivité. Il suffit d'avoir présent à l'esprit qu'une analyse nécessite diverses sortes d'interprétations, selon le type de son, et surtout selon les interrogations qui motivent l'analyse. Il convient donc de s'approprier les analyses, d'apprendre à les lire et aussi d'en connaître les limites. On parviendra ainsi à choisir rapidement le format d'analyse adapté à la question posée et, peu à peu, à éviter les erreurs d'interprétation. Plusieurs logiciels libres d'accès sont disponibles. Citons principalement Praat et Wavesurfer (voir l'annexe F, rubrique Logiciels). Ces logiciels peuvent fournir une foule de renseignements pour peu que l'on soit un minimum familier des principes de l'analyse des paramètres du son.

1.3. Les étapes d'une analyse

1.3.1. La saisie des vibrations

Pour commencer, il faut pouvoir capter le son au moyen de dispositifs qui transforment les vibrations en variations de tension électrique (figure 2.2). Les capteurs les plus usités sont sensibles aux variations locales de la pression ou de la vitesse de l'air (microphones), aux variations d'un champ magnétique (micro de guitare), aux déplacements d'un corps solide (accéléromètres). Le signal issu du capteur,

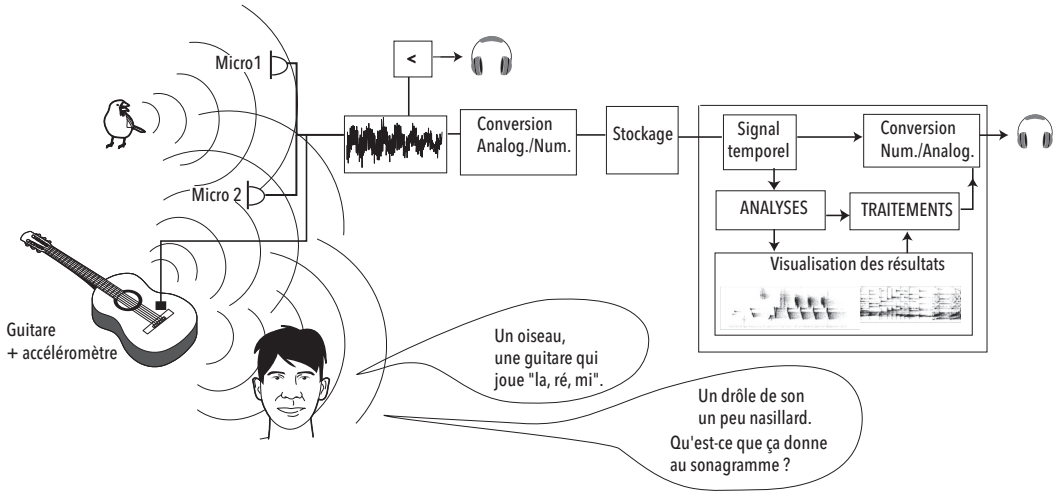


Figure 2.2 Étapes d'une analyse acoustique.

amplifié, est traité différemment selon les appareils enregistreurs. Aujourd'hui, ce signal continu est numérisé à l'aide d'une carte dite analogique-numérique, puis stocké sur un support numérique : ordinateur, mémoire de masse.

L'acquisition numérique garantit une grande fidélité de l'enregistrement et permet aussi de s'affranchir des problèmes mécaniques inhérents aux appareils analogiques qui affectaient souvent la régularité de défilement et donc la reproduction des fréquences (magnétophones, cassettes). En revanche, l'amplitude relative des fréquences reste toujours très dépendante des qualités du capteur et surtout de sa position dans le champ acoustique. Le contrôle de l'enregistrement à l'aide d'un casque permet d'adapter au mieux la captation pour l'analyse que l'on se propose d'interpréter. (Voir Fromentel, 2010, *La technologie audionumérique*)

1.3.2. La restitution : écouter les sons que l'on analyse

Pour tirer le meilleur parti des analyses acoustiques, en particulier celle des sons musicaux, il faut pouvoir écouter à tout instant les sons que l'on est en train d'analyser. L'écoute au casque est pratique, car elle a pour avantage d'isoler des bruits extérieurs. En revanche, elle prive l'auditeur de la possibilité d'explorer le champ sonore par les petits mouvements de la tête qui lui fournissent de fines sensations différentielles entre les deux oreilles. Cette exploration est particulièrement précieuse pour l'écoute des qualités spectrales et des accords complexes. L'écoute en champ libre nécessite une chaîne d'amplification et de bonnes enceintes de reproduction – la sortie aérienne d'un ordinateur portable n'étant utile que pour contrôler l'identité du fichier et le bon fonctionnement de la connectique.

1.3.3. Pour une analyse active : les transformations et la resynthèse

Quand nous entreprenons une analyse, c'est pour trouver des réponses aux interrogations que nous formulons. Cependant il y a un tel écart entre les grandeurs mesurables d'un signal sonore et nos stratégies d'écoute, en particulier les indices que nous utilisons pour discriminer les sons, que l'interprétation est souvent problématique. Pour progresser, il ne faut pas hésiter à manipuler le son de diverses façons. La plupart des logiciels permettent de faire une sélection dans le son, de mettre un fragment en boucle pour l'écouter isolément, d'opérer des transpositions

fréquentielles ou temporelles, d'inverser le temps (lecture à rebours dite *reverse)*, ou encore de faire des montages pour écouter deux sons à comparer dans une succession rapide. Quelques logiciels offrent des transformations plus avancées que l'on peut effectuer directement sur la représentation sonagraphique. Toute technique de travail sur le son qui permet d'associer en temps réel les transformations du son, l'écoute du résultat et la visualisation est extrêmement puissante. L'écoute analytique requise pour tirer parti des analyses est fort différente de l'écoute musicale habituelle. Elle s'éduque et s'affine par la pratique.

Le travail se déroule couramment selon trois étapes. Étape 1 : première visualisation d'ensemble par le sonagramme et travail à l'oreille pour repérer les éléments pertinents (zones de fréquence, bruit, transitoires). Étape 2 : production d'analyses ciblant les points pertinents. Étape 3 : vérification par la resynthèse en transformant les paramètres que l'on a repérés comme étant responsables des effets perceptifs.

2. Première exploration : une « mise en bouche »

Son 2.1 (5")

Quatre sons de « bouche »

Voici quatre sortes de sons que chacun peut produire avec sa bouche et son larynx : un sifflement vibré, un chuintement, un /a/ chanté avec une intonation ascendante et deux claquements de langues (Son **2.1**). La figure 2.3 en montre l'image sonagraphique[3], image avec laquelle nous sommes déjà familiarisés.

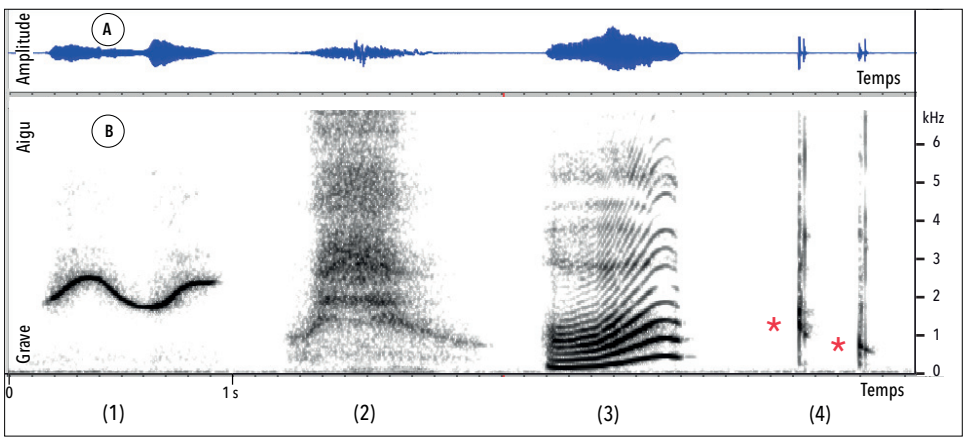


Figure 2.3 Analyse sonagraphique de quatre sons « buccaux ». (1) Sifflet oral. (2) Chuintement « chouiiiou ». (3) /a/ chanté sur une intonation ascendante. (4) Deux claquements de langue. Zone A : tracé du signal sonore (amplitude/temps). Zone B : sonagramme. (Analyse Audiosculpt)

Du point de vue perceptif, cette représentation possède deux avantages décisifs : l'importance donnée aux variations temporelles, donc aux « allures » dynamiques et fréquentielles des événements sonores, et le fait que les trois dimensions du son – fréquence, intensité, temps – soient combinées dans une image globale prégnante, obtenue aujourd'hui par calcul de la transformée de Fourier. (Nous verrons plus en détail cette représentation au § 3.3, page 43.)

3. Le terme sonagramme (initialement réservé aux documents produits par l'appareil de la marque Sona-Graph) est aujourd'hui passé dans le langage courant et désigne les représentations du même type, quel qu'en soit la technique de production. Les premières réalisations ont eu lieu aux États-Unis pendant les années 1940 (voir Koenig, Dunn & Lacy, 1946, *The sound spectrograph*). Dès son introduction en France en 1956, les spécialistes de la parole et de l'acoustique animale (oiseaux) s'en emparent. À la faculté des sciences de Paris, E. Leipp et M. Castellengo l'utilisent pour l'analyse des bruits et des sons musicaux ainsi que pour l'étude des musiques traditionnelles (voir les *bulletins du GAM* n°s 4, 6, 8, 12, et suivants).

Ce type de document associe généralement deux représentations : le signal temporel (A) et le sonagramme (B).

Sur le sonagramme, le temps est représenté sur l'axe horizontal et le contenu fréquentiel sur l'axe vertical ; le degré de noircissement du tracé est lié à l'intensité. La lecture de ce document peut s'effectuer de différentes façons.

Forme globale. Tout d'abord on remarque que le tracé permet de différencier globalement la texture sonore et l'évolution temporelle de ces quatre signaux. Le sifflet oral (1) est représenté par un trait ondulant, le « chouiiiou » (2) par un nuage de points, la voyelle chantée (3) par une série de traits superposés d'écartement variable, les claquements de langue (4) par deux groupes de traits verticaux.

Allures fréquentielles. Les évolutions dans le sens vertical sont liées à la sensation de hauteur du grave à l'aigu, de façon analogue à une partition musicale. Si nous prêtons attention aux variations de hauteur sonore de ces quatre signaux, nous voyons que l'ondulation du tracé (1) représente la mélodie de la phrase sifflée. Dans l'émission chuintée, le mouvement ascendant/descendant est visible sur la partie inférieure du tracé. L'intonation de la voyelle (3) est figurée par le mouvement ascendant des raies. Des deux claquements de langue (4), le second est perçu plus grave, ce qu'indique le déplacement de la zone noire inférieure signalée par les astérisques.

Allures dynamiques. Elles se présentent sous deux formes : globales et locales. Sur la figure 2.4, la courbe d'amplitude A rend compte des variations *globales* du son en fonction du temps : le début du son ou transitoire d'attaque, les variations pendant la tenue, la décroissance et l'extinction. Prenons l'exemple du sifflet (1). La variation de fréquence visible sur le sonagramme est lente et continue et, pourtant, nous entendons deux *notes* : un son aigu et un glissé à la quarte inférieure. Le fait que l'amplitude présente deux maxima associés au son aigu (mib_6) renforce l'impression de stabilité de cette note. Pour le claquement de langue (4), la courbe A montre de façon très nette une double impulsion caractéristique.

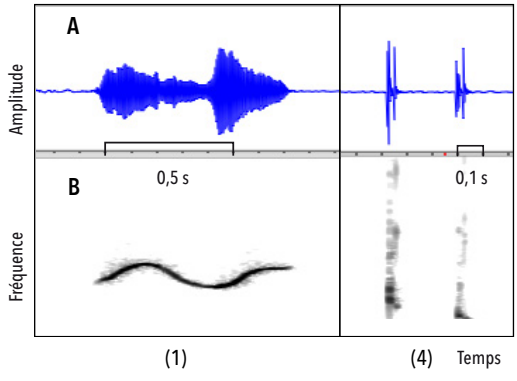


Figure 2.4 Allures des variations d'amplitude d'un son continu (sifflet) et d'un son impulsif (claquement de langue).

Sur le sonagramme (B), les variations *locales* d'intensité sonore sont indiquées par le noircissement et la largeur du tracé. Le réglage des seuils d'affichage des points blancs (intensités faibles) et noirs (intensités fortes) permet de régler à la fois la

luminosité et le contraste des images. On peut choisir (figure 2.5) entre une image exhaustive mais grise (B1) ou une image contrastée mais plus schématique (B2). L'emploi d'une charte de couleurs associée à l'échelle des intensités, variable selon les logiciels, permet de jouer sur les seuils intermédiaires tout en conservant une dynamique globale satisfaisante (B3 et B4). Dans notre exemple l'intensité croît du noir au blanc en passant par le bleu, le jaune et le rouge.

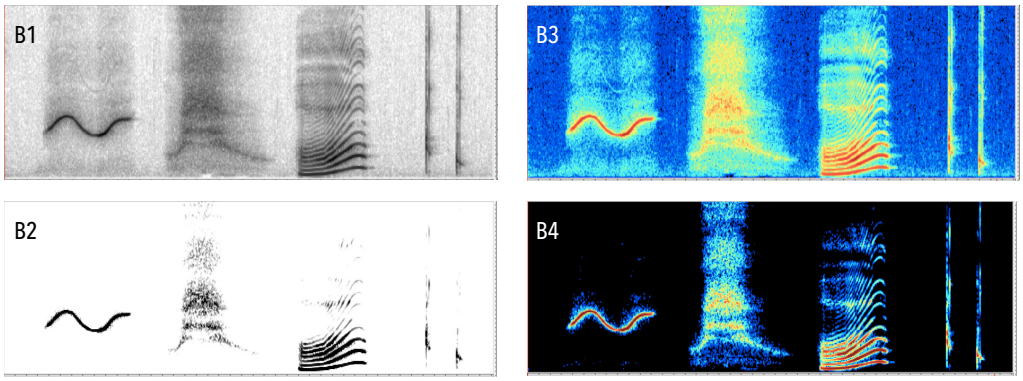


Figure 2.5 Véritable photographie du son, l'image sonagraphique gagne a être travaillée en jouant sur les contrastes et, au besoin, sur la couleur, afin de mettre en valeur les éléments pertinents du point de vue perceptif. (B1) faible contraste (dynamique 20-120). (B2) fort contraste (dynamique 31-66). (B3) couleurs *Hot-Cold* (dynamique 20-120). (B4) couleurs *Hot-Cold* (dynamique 32-82). (Analyses Audiosculpt)

La représentation spectrographique constitue une bonne introduction à toute analyse des sons. Son caractère intuitif permet de repérer assez vite les éléments saillants d'une séquence sonore, de délimiter les zones de fréquence à explorer et de discriminer les événements pertinents par rapport aux bruits de fond et aux parasites éventuels. Cependant, pour accéder à des données mesurables, il faut laisser de côté l'une des trois dimensions du son.

3. Les paramètres du signal sonore

3.1. La temporalité et l'analyse des vibrations sonores

Le problème crucial de l'analyse acoustique est celui de la décomposition de l'onde (ou du mouvement) en ondes élémentaires avec leurs fréquences et leurs amplitudes respectives. Or, la notion de fréquence se définissant dans le temps, il devient rapidement évident que l'analyse précise des composantes avec leur fréquence et leur amplitude implique de figer le temps. Dans son ouvrage, Helmholtz prend bien soin de préciser que l'analyse harmonique dont il rend compte ne concerne que la partie stable des sons, « en faisant abstraction (des) particularités relatives au commencement et à la fin du phénomène sonore » (Helmholtz, 1874, chapitre V, page 94).

Ainsi, pendant plus d'un siècle les sons périodiques, dits « musicaux », sont restés les parangons de l'analyse harmonique. Nécessairement stables par définition (car, en toute rigueur, un son périodique n'a ni commencement ni fin), ils se prêtent bien à l'analyse de Fourier. Le tableau de la figure 2.6 résume les principales étapes historiques de l'analyse acoustique. Ce n'est que récemment que les techniques

d'analyse spectrale fondées sur le filtrage ont ouvert l'accès à l'analyse spectrale de sons non périodiques et de bruits, pour peu que l'on puisse trouver le meilleur compromis fréquence-temps.

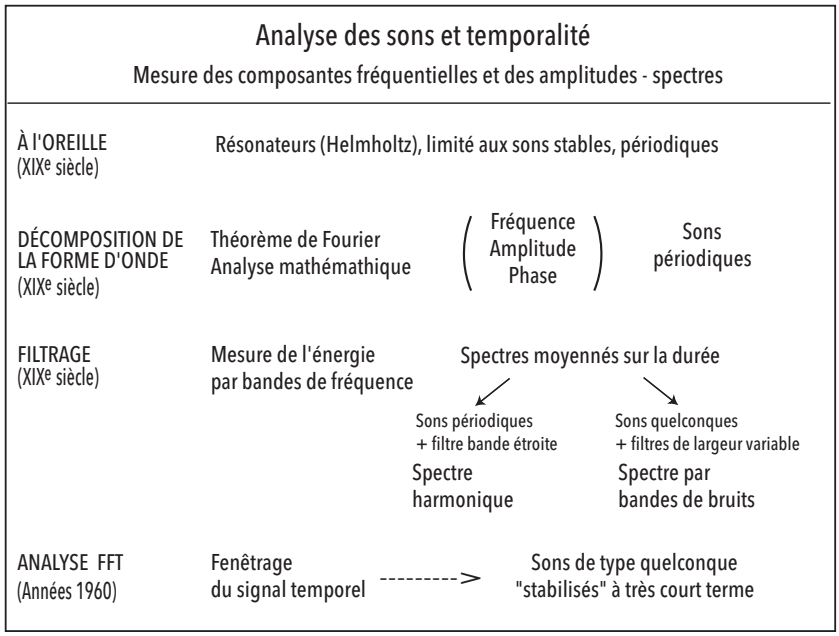


Figure 2.6 Jalons des progrès de l'analyse acoustique des signaux sonores pour saisir les variations temporelles des signaux.

3.2. La mesure de la fréquence : un problème difficile[4]

La notion de fréquence est liée au temps : c'est le nombre d'oscillations effectuées dans une unité de temps (seconde). Plus le son est stable sur une longue durée, plus on aura de chances de mesurer la fréquence avec précision. Or, les sons intéressants, captivants pour l'oreille, sont justement ceux qui varient dans le temps, de façon subtile et complexe.

Nous avons souligné à plusieurs reprises l'intérêt des représentations de type sonagraphique qui fournissent l'évolution *temporelle* des fréquences et des amplitudes[5]. L'usage de la transformée rapide de Fourier à court terme (ou STFT[6]) aurait-il miraculeusement résolu ce dilemme ? Oui et non. Oui, car il est maintenant possible d'analyser tout type de son (sons inharmoniques, bruits) et d'accéder aux phénomènes très brefs (impulsions, transitoires). Non, car à chaque analyse il faut choisir de privilégier soit l'analyse fine de la fréquence, soit le suivi des petites variations temporelles : le couple antagoniste temps-fréquence est incontournable.

Les musiciens qui ont un grand entraînement pour écouter la justesse des sons, et en particulier ceux qui ont l'oreille absolue, s'étonnent souvent qu'il n'existe pas

4. Wikipedia classe les méthodes de mesure de la fréquence (*Pitch Detection Algorithms* ou PDA) en méthodes temporelles et méthodes fréquentielles.
5. Les premières analyses sonagraphiques étaient fondées sur une technique de filtrage glissant effectué 144 fois entre 0 et 6000 Hz, ce qui permettait de contourner le problème au prix d'une absence de précision, sans compter qu'il fallait près de 6 min pour analyser 2,4 s de son !
6. L'analyse des composantes d'un son périodique par transformée de Fourier est longue et complexe. L'usage d'algorithmes a permis de gagner en rapidité : c'est la transformée rapide de Fourier ou FFT *(Fast Fourier Transform)*. Quand elle est appliquée à un fragment de son très bref, le calcul porte le nom de transformée de Fourier à court terme *(Short-Time Fourier Transform)*. Voir § 3.3, page 43 et Glossaire.

encore de méthode simple et totalement fiable pour mesurer et transcrire automatiquement les mélodies qu'ils entendent. Il faut tout d'abord rappeler que la sensation de hauteur perçue fait appel à plusieurs stratégies cognitives différentes, adaptées à la diversité acoustique des sons réels instrumentaux – stratégies que nous sommes loin de comprendre complètement et que nous étudions en détail au chapitre 3, § 2.9 et tout au long du chapitre 6. Quant à l'extraction d'une voix dans une polyphonie instrumentale, telle qu'un musicien peut le faire à l'oreille, elle n'est pas encore réalisable automatiquement.

En pratique, les méthodes de mesure diffèrent selon que les signaux sont relativement stables ou affectés de variations très rapides.

3.2.1. Une estimation manuelle de la fréquence des sons périodiques

Il suffit de zoomer ou d'agrandir l'échelle temporelle. Voici un exemple pris sur l'analyse du Son **2.1**.

En partant de la courbe du signal temporel (voir courbe A de la figure 2.7), prélevons de petites portions du signal (traits rouges sur le tracé) dont la représentation dilatée temporellement apparaît au-dessous (zone A'). Sur cette nouvelle échelle (intervalle de 15 ms), la succession des alternances positives et négatives du signal sonore devient visible. Les formes d'onde du chuintement (2) et du claquement (4) sont erratiques. Celles du sifflet (1) et de la voix (3) présentent des alternances qui se répètent de façon régulière : elles sont quasi périodiques et nous allons pouvoir en mesurer la fréquence fondamentale.

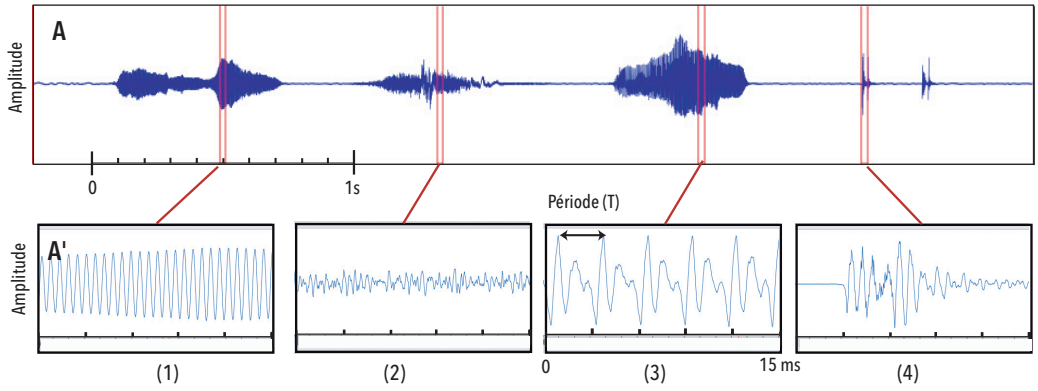


Figure 2.7 Analyse du Son 2.1. Zone A : forme d'onde du signal sur la durée totale (4 s). Zone A' : forme d'onde dilatée (15 ms) de la partie sélectionnée en rouge sur la zone A. Sifflet (1), chuintement (2), voix (3), premier claquement de langue (4). Seuls les sons (1) et (3), quasi périodiques, montrent une forme d'onde régulière. (Analyse Audiosculpt)

Il faut tout d'abord repérer la période (T), mais, comme celle-ci est souvent d'une durée brève et que la précision fait défaut, on contourne la difficulté en comptant plusieurs périodes. Dans l'exemple (3), la durée de cinq périodes est de 14 ms, donc T (ms) = 2,8 ms, et f (Hz) = 1/(2,8) 10^{-3} soit environ f = 350 Hz, ce qui est un fa_3. La mesure n'est pas précise, car le son varie rapidement (glissando), mais l'opération que nous venons de faire fournit déjà une estimation locale, un ordre de grandeur. En répétant l'opération au tout début et à la fin du glissando, on peut estimer l'intervalle de variation du glissando à environ une octave (la_2-la_3).

3.2.2. La méthode temporelle du suivi de la fréquence fondamentale

Une des premières méthodes d'extraction automatique de la fréquence fondamentale, mise en œuvre pour le tracé de l'intonation de la voix parlée, a consisté à filtrer la fréquence fondamentale. Sur le signal sinusoïdal ainsi obtenu, il est aisé de repérer les passages par zéro de la forme d'onde, ce qui permet ensuite de mesurer la période pour remonter à la fréquence. Cette méthode n'est utilisable qu'avec des sons possédant toujours de l'énergie dans le premier harmonique et ne fonctionne bien de façon automatique que dans l'ambitus d'une octave. Avec le développement des techniques de traitement du signal sont apparus différents algorithmes de calcul.

3.2.3. Les algorithmes d'extraction automatique de la fréquence fondamentale par autocorrélation

Il s'agit de techniques applicables à des signaux complexes périodiques. Le fait de ne prendre en compte qu'un petit nombre d'oscillations en fait des méthodes assez bien adaptées aux signaux de fréquence variable mais localement harmoniques.

En termes simples, l'autocorrélation consiste à comparer le signal à lui-même avec un décalage temporel variable. Si le signal est périodique, l'algorithme détecte un maximum de coïncidences lorsque le décalage temporel atteint une période, ce qui est la réponse cherchée. Le résultat n'est conservé que lorsque la similitude se confirme sur deux à trois périodes consécutives. Cette méthode est utilisée par le logiciel Praat (voir les références à l'annexe F, rubrique Logiciels).

Avant de lancer un calcul de la fréquence fondamentale, il est nécessaire de spécifier les limites inférieure et supérieure de la variation en fréquence à explorer pour éviter des erreurs d'octave ou de « fausses » détections.

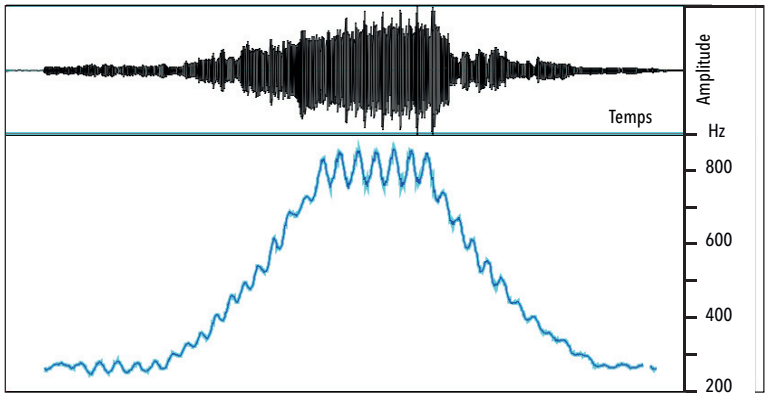


Figure 2.8 Analyse de la fréquence fondamentale d'une gamme diatonique chantée. Voir chapitre 6, § 4.3.3 pour l'exemple sonore et les commentaires. Sur la partie supérieure : amplitude en fonction du temps ; sur la partie inférieure : tracé de la fréquence (de 200 à 900 Hz) en fonction du temps (6,25 s). (Analyse Praat)

Son 2.2 (7")

Gamme chantée avec vibrato

La figure 2.8 représente l'analyse d'une gamme chantée (Son **2.2**) de do_3 (260 Hz) à sol_4 (785 Hz), par mouvement ascendant et descendant. Pour le calcul de la fréquence fondamentale, nous avons spécifié un intervalle plus large : de 200 à

900 Hz. Selon les besoins, le résultat peut s'afficher en échelle linéaire ou logarithmique, par exemple en demi-tons (voir page 53).

3.3. La transformée de Fourier à court terme et la dualité temps/fréquence

La transformée rapide de Fourier à court terme, ou STFT[7], permet d'effectuer l'analyse de signaux quelconques : sons apériodiques, bruits continus, impacts et, bien sûr, les sons périodiques. Le principe consiste à prélever une fraction temporelle du signal, ou « fenêtre » temporelle, à l'intérieur de laquelle est calculée la répartition de l'énergie en fonction de la fréquence, c'est-à-dire le spectre. L'opération se répète pour la fenêtre suivante, et ainsi de suite jusqu'à la fin du son (figure 2.9).

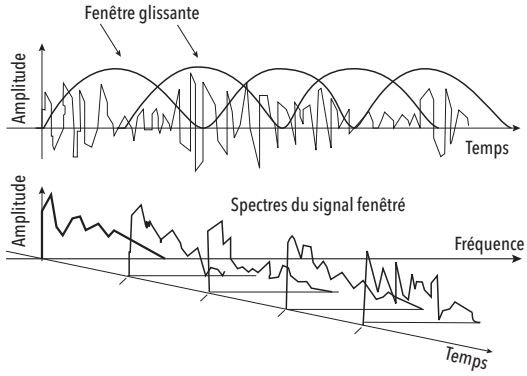


Figure 2.9 Principe des fenêtres glissantes de la transformée de Fourier.

Battier, M., manuel Audiosculpt v. 1.2ß1, Ircam, Centre Pompidou.

En toute rigueur, il faut définir la forme de la fenêtre, sa durée temporelle, Δt, et le « pas » du décalage entre deux fenêtres successives, car celles-ci se recouvrent. En pratique, les choix du type de fenêtre et du pas d'avancement sont proposés par défaut et il faut surtout se préoccuper de bien définir la largeur (ou durée) de la fenêtre, ce qui nous ramène à la relation temps/fréquence mentionnée plus haut.

On peut traduire cette relation par l'expression $\Delta t \times \Delta f = \text{constante}$, dans laquelle le signe Δ représente l'intervalle de définition (précision) de la mesure. Le produit des deux termes étant constant, lorsque l'un augmente, l'autre diminue : c'est une relation d'incertitude. Dans le plan fréquence/temps, un « quanta » de son peut être représenté par un rectangle (voir figure 2.10). Pour Δt petit (t_1), le rectangle est vertical ; sa position est précise en temps mais couvre une large bande de fréquence. Pour Δt grand (t_2), le rectangle est horizontal ; il est précis en fréquence mais flou temporellement.

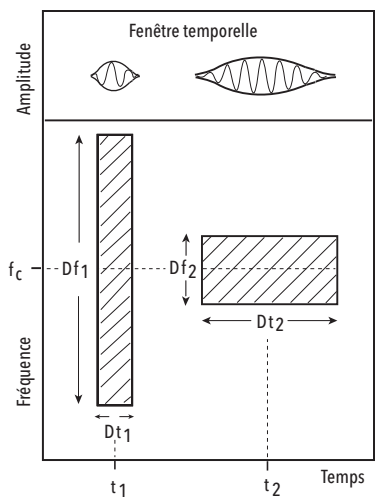


Figure 2.10 Diagramme de Gabor d'après H. Pollard. Analyse d'un « quanta » de son avec deux fenêtres temporelles de durées différentes (voir bandeau supérieur) : courte (Δ t1) et longue (Δ t2).

D'après Pollard, H., 1988, *Acustica*, **65**, figure 7b,

7. Voir note 6 et le glossaire.

Dans l'intervalle Δt de la fenêtre, les variations du signal en amplitude ou en fréquence ne sont pas prises en compte. Si on choisit une fenêtre de durée courte pour suivre les variations temporelles, la résolution en fréquence devient mauvaise. En pratique, il faut soit s'en tenir à un compromis, soit effectuer deux analyses complémentaires avec des fenêtres de durées différentes. Le néophyte est souvent désorienté par le fait que la terminologie utilisée et les caractéristiques de la fenêtre (en ms, ou en nombre d'échantillons) varient selon les logiciels. Reportez-vous au § 4.13, page 71 pour une comparaison de trois logiciels différents.

3.3.1. Un exemple d'analyse : l'impulsion du cor des Alpes

Son 2.3 (2")

Le « pop » du tuyau du cor des Alpes

Voici une illustration simple et spectaculaire de l'incidence du choix de la fenêtre d'analyse sur la représentation d'un son (figure 2.11). Nous avons retenu l'exemple déjà présenté chapitre 1, page 13 : le son obtenu par la brève excitation du tuyau d'un cor des Alpes (Son **2.3**). C'est un son grave, $f = 46{,}72$ Hz ($fa\#_0$), T = 0,0214 s.

La largeur de la fenêtre d'analyse contient un nombre entier d'échantillons (ou points). Le choix de sa durée (Δt en secondes), a pour corrélat une résolution donnée de la fréquence (Δf).

L'analyse B est faite avec une fenêtre d'analyse étroite : Δt = 0,011 s, valeur inférieure à la période du son, ce qui permet de séparer distinctement les impulsions successives se produisant à chaque aller et retour dans le tuyau. En contrepartie, la résolution en fréquence n'est que de 86 Hz : le spectre est continu.

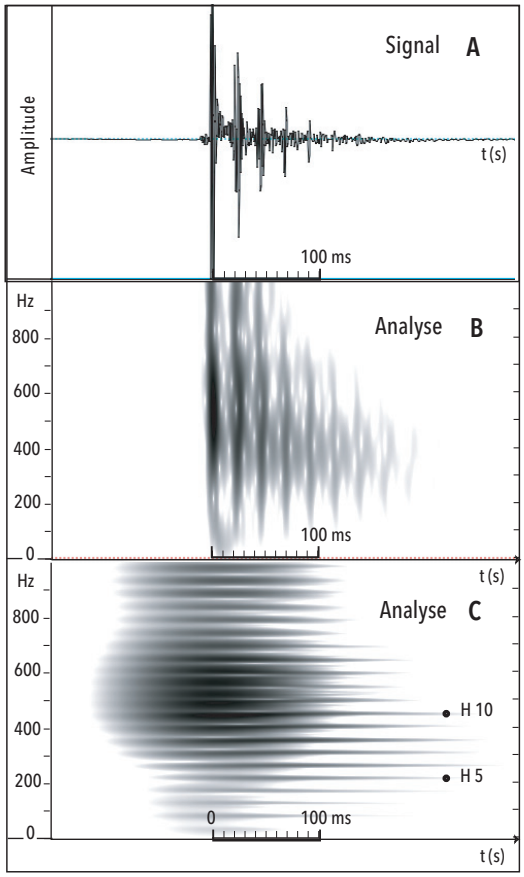


Figure 2.11 Analyse du train d'impulsions recueilli au pavillon d'un cor des Alpes (chapitre 1, § 2.3.2). En A, signal temporel. En B et C, sonagrammes réalisés avec deux fenêtres temporelles différentes. Pour B : Δt = 0,011 (Δf = 86 Hz). Pour C : Δt = 0,185 s (Δf = 5,35 Hz) ; (Logiciel Praat). La fréquence d'échantillonnage du son est 44,1 kHz.

L'analyse C est faite avec une fenêtre large : Δt = 0,185 s et une résolution en fréquence de 5,35 Hz. Cette fois l'image montre avec précision un spectre de raies harmoniques et la répartition d'énergie sur chacun d'eux. En contrepartie, les transitoires (début et fin du son) sont très mal définis.

Les deux analyses fournissent deux informations complémentaires – l'une temporelle, l'autre fréquentielle – du même phénomène sonore. Selon la taille de la

fenêtre choisie pour faire l'analyse, la représentation visuelle d'un son peut donc changer d'aspect de façon assez considérable.

3.3.2. L'analyse d'un son de tuba

L'exemple suivant, similaire, est l'analyse d'un son grave entretenu joué au tuba (Son **2.4**), un do_1 dont la fréquence fondamentale est $f = 66$ Hz, et la période $T = 1/66 = 0{,}015$ s (figure 2.12). L'analyse A est effectuée avec une FFT de 256 échantillons (*samples),* celle de B avec 4096 échantillons.

Son 2.4 (3")

Son de tuba (do_1)

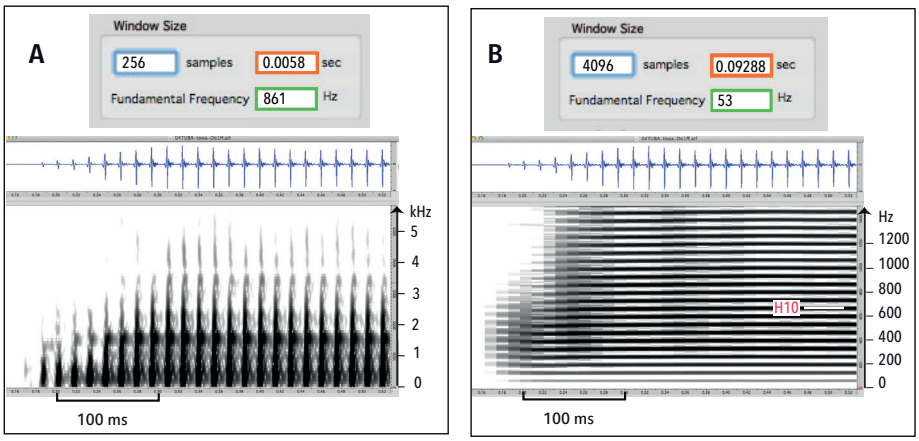


Figure 2.12 Analyses FFT d'un son de tuba : note do_1, environ 66 Hz, avec deux fenêtres temporelles différentes. De haut en bas : paramètres de l'analyse ; signal d'amplitude ; sonagramme. En A, $\Delta t = 5{,}8$ ms. La succession des impulsions (visibles sur l'amplitude) s'affiche nettement sur le sonagramme. En B, $\Delta t = 92{,}8$ ms. Sur le sonagramme, les impulsions ont disparu au profit d'un spectre de raies. (Logiciel Audiosculpt)

En A, la fenêtre temporelle est de petite taille ($\Delta t = 5{,}8$ ms). L'analyse spectrale (sonagramme) permet de suivre les événements temporels du transitoire initial du son.

En B, avec une fenêtre de plus grande taille ($\Delta t = 92{,}8$ ms), l'image sonagraphique de l'établissement du son est floue, mais la résolution en fréquence est excellente. Le spectre harmonique résolu apparaît alors sous la forme habituelle de raies parallèles. L'agrandissement de l'échelle de fréquence permet d'évaluer la fréquence du 10e harmonique situé entre 600 et 700 Hz, ce qui est en accord avec la valeur 66 Hz du fondamental. Remarquez que l'on peut aussi compter le nombre d'impulsions sur la première analyse, environ 6,5 dans l'intervalle de 100 ms (soit 65 dans 1 s) ce qui est une autre façon d'accéder à la fréquence fondamentale.

Sur le panneau d'affichage des paramètres, *fundamental frequency* (53 Hz) indique la limite inférieure en fréquence de l'extraction automatique de la fréquence fondamentale par calcul des FFT proposée dans le logiciel (pour une explication plus détaillée voir le tableau 3, page 73). Comme pour l'exemple du paragraphe précédent, on remarque que l'analyse par FFT offre une double représentation de la fréquence fondamentale d'un son périodique : sur l'axe temporel (nombre d'impulsions dans une seconde), et sur l'axe des fréquences (spectre harmonique).

Terminologie

La terminologie utilisée en analyse acoustique est loin d'être unifiée. Des termes équivalents mais différents circulent parmi les experts : métrologie acoustique, traitement du signal, spécialistes de la parole. À cela s'ajoute l'usage des termes anglais des logiciels, eux aussi divers, qui ne sont pas toujours équivalents aux termes français.

Les représentations de l'évolution de la fréquence et de l'amplitude en fonction du temps (f/I/t) sont désignées tantôt par sonagramme, *sonogram* ou encore spectrogramme. Ce dernier terme ne doit pas être confondu avec le spectre ou *spectrum section* qui est strictement le plan fréquence/intensité, pris à un instant (ou dans une durée de temps) défini au préalable

Le plan fréquence/temps sans indication d'intensité est le plan mélodique par référence à l'écriture musicale. On y représente l'évolution de la fréquence fondamentale des sons périodiques, ce qui correspond à l'intonation de la voix parlée ou à la hauteur musicale d'une mélodie. En traitement du signal, la fréquence fondamentale extraite par calcul est appelée le *pitch* ou la f_0 ce qui n'est pas sans inconvénient. En effet, le terme *pitch* désigne, en psychoacoustique, la fréquence perçue, souvent différente de la fréquence mesurée. Par ailleurs f_0, qui est la fréquence fondamentale, est aussi l'harmonique 1. On voit que l'usage de ces termes doit être pratiqué en toute lucidité. En ce qui concerne l'intensité sonore, on dispose de deux représentations : l'enveloppe temporelle du signal, généralement en échelle linéaire, et la courbe d'intensité en fonction du temps, en dB, unité logarithmique qui rend compte de l'évolution dynamique.

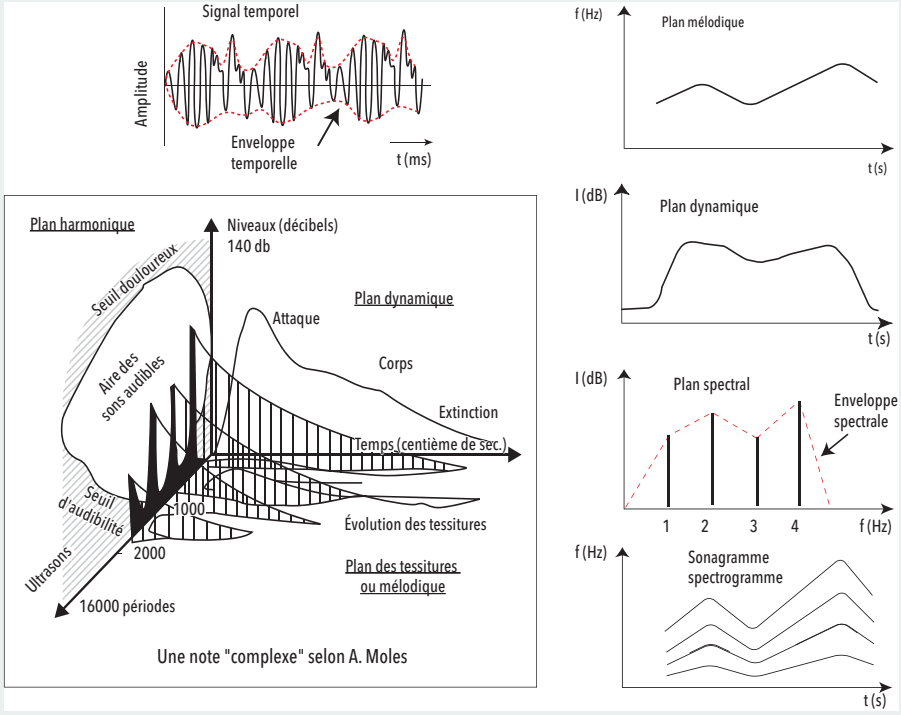


Figure 2.13 Représentations du son et terminologie.
Source de la figure encadrée : Moles, A., 1952, p. 431.

3.4. L'amplitude globale : l'enveloppe temporelle, la dynamique

En un point donné de l'espace, un microphone recueille la somme des amplitudes des ondes acoustiques qui s'y croisent. La trace la plus immédiate en est la courbe de la variation d'amplitude du signal en fonction du temps, dont nous avons déjà vu l'intérêt pour l'interprétation des analyses spectrographiques (figures 2.1 et 2.4). L'amplitude du signal temporel est proportionnelle à l'énergie produite, mais, si l'on veut s'approcher au mieux de la « sensation d'intensité » perçue, il vaut mieux utiliser une unité logarithmique : le décibel (voir chapitre 3, § 2.4, la loi de Fechner et tout le chapitre 5). Pour une présentation comparée des échelles linéaire et logarithmique, voir Glossaire, page 521.

3.4.1. L'évolution temporelle d'un son (attaque, tenue, extinction)

Voici deux exemples (Son **2.5**) : un vibraphone et un tuyau d'orgue jouant la même note do_4 (figure 2.14).

Son 2.5 (6")

Note do_4 jouée par un vibraphone et par un tuyau d'orgue

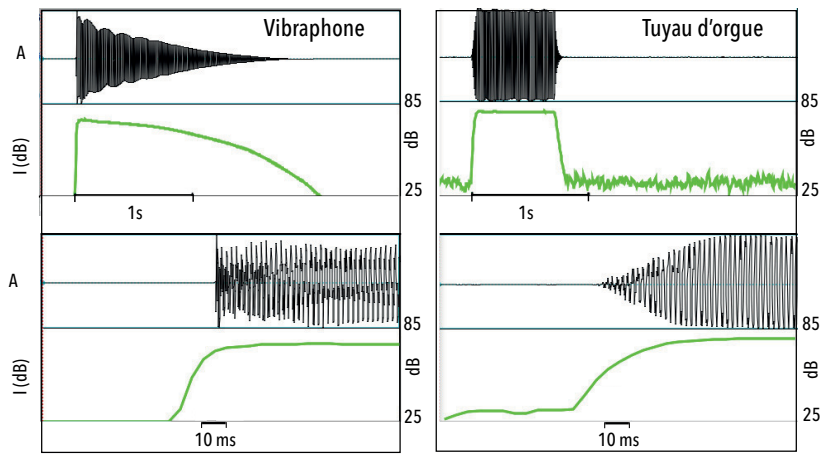


Figure 2.14 Comparaison de l'enveloppe temporelle de l'amplitude du signal sonore (noir) et de l'intensité calculée en dB (vert). Les analyses comparées des deux signaux sont à la même échelle temporelle (3 s). En bas, mêmes analyses zoomées sur le transitoire d'attaque. (Logiciel Praat)

L'amplitude du son percuté croît très rapidement, atteint son maximum au moment de l'attaque, puis décroît régulièrement. Pour le son du tuyau d'orgue, le début et la fin sont similaires et l'amplitude reste quasi constante pendant la durée de l'entretien. Les courbes tracées en vert sont celles de l'intensité calculée en dB. On remarque que le profil de la courbe d'extinction du vibraphone en dB a changé de forme. Pour certains auditeurs, cette courbe arrondie rend mieux compte de la sensation de tenue que procure le vibraphone. Nous avons choisi intentionnellement de régler de façon identique, pour les deux sons, les valeurs minimale et maximale de l'intensité, respectivement à 25 dB et 85 dB, soit une dynamique totale de 60 dB. Les ondulations de la courbe inférieure du tuyau d'orgue sont dues au bruit de la soufflerie.

Dans la partie inférieure de la figure 2.14, l'agrandissement de l'échelle temporelle permet de focaliser l'analyse sur le transitoire d'attaque. On voit sur la courbe verte

un artefact imputable au calcul de l'intensité. D'une part les courbes en dB commencent avant le début réel du son, d'autre part, celle du vibraphone est à peine moins inclinée que celle du tuyau d'orgue. La durée réelle d'établissement du son est lisible sur la courbe d'amplitude A (courbe noire).

L'enveloppe temporelle (ou dynamique) est l'un des paramètres de base des systèmes de synthèse sonore qui invitent à modifier séparément le transitoire d'attaque, la tenue et l'extinction des sons. Toutefois, l'enveloppe d'intensité ne suffit pas à caractériser un transitoire. Il se produit souvent des événements fréquentiels de faible énergie mais très saillants perceptivement, qu'on ne peut déceler que sur un sonagramme. Voir aussi § 3.6, page 51.

3.4.2. Les variations d'amplitude et le rayonnement des modes vibratoires

Voici deux exemples de variation de l'amplitude sonore dus aux modifications du champ sonore (Son **2.6**).

Verre frotté. Lorsqu'on fait chanter un verre en le frottant avec un doigt mouillé, le bord se déforme en présentant une succession de zones ventrales et nodales (voir le Glossaire page 528). Celles-ci se déplacent pendant que le doigt tourne en produisant les variations d'amplitude globale visibles sur la courbe de la figure 2.15.

Son 2.6 (21")

Verre puis petit gong

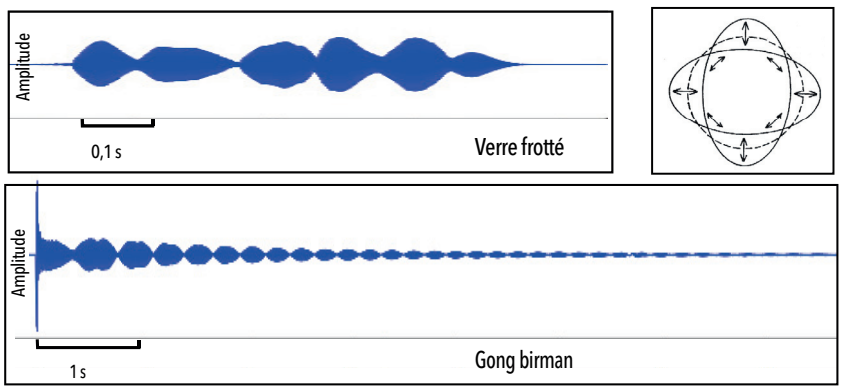


Figure 2.15 Deux exemples de variations d'amplitude. Partie supérieure : le verre frotté déjà présenté chapitre 1, page 9. À droite : déformations du bord supérieur du verre. Partie inférieure : analyse d'un petit gong birman.

Petit gong Birman. Le petit gong de la figure 2.16 est une plaque de laiton suspendue par un cordonnet que l'on tord avant de frapper la plaque. Au cours de l'extinction, le fil se déroule et le gong tourne sur lui-même en produisant d'importantes variations d'amplitude dues aux changements de direction du rayonnement du mode principal de la plaque (voir figure 2.15).

Les variations de l'amplitude perçues de ces deux exemples n'ont rien de commun avec le phénomène connu sous le nom de battement (Glossaire page 516).

Figure 2.16 Gong birman.
Photo M. C.

3.5. Les spectres : spectres à court terme et spectres moyennés

Un spectre représente la répartition de l'amplitude en fonction des fréquences. Fondamentalement, sa réalisation nécessite un filtrage associé à une mesure de l'amplitude. Il existe un grand nombre de méthodes et une très grande diversité de représentations spectrales selon les types de filtrages choisis, selon la durée pendant laquelle se fait le calcul, et selon les modalités de représentation des résultats, en particulier pour le choix de l'échelle de fréquence qui peut être linéaire ou logarithmique (Glossaire, page 521).

3.5.1. Les spectres à court terme

La représentation sonagraphique est obtenue par juxtaposition d'une multitude de spectres successifs se recouvrant partiellement. On a vu que son intérêt réside dans l'affichage des variations temporelles, au prix d'une certaine imprécision visuelle de l'amplitude et de la fréquence. Les informations ont été calculées et peuvent être représentées différemment sur les deux axes fréquence/intensité. Bien que le temps soit exclu de la représentation spectrale, ce paramètre est à prendre en compte au moment de la requête d'un spectre, puisqu'il faut bien définir la tranche temporelle à l'intérieur de laquelle s'effectue le calcul. Comme on le sait, celle-ci conditionnera la résolution fréquentielle d'un spectre dit « instantané » ou à court terme[8]. La position du curseur et la résolution en fréquence doivent être choisies en fonction des besoins. La figure 2.17 montre quatre spectres pris à un instant donné des sons présentés figure 2.3, page 37.

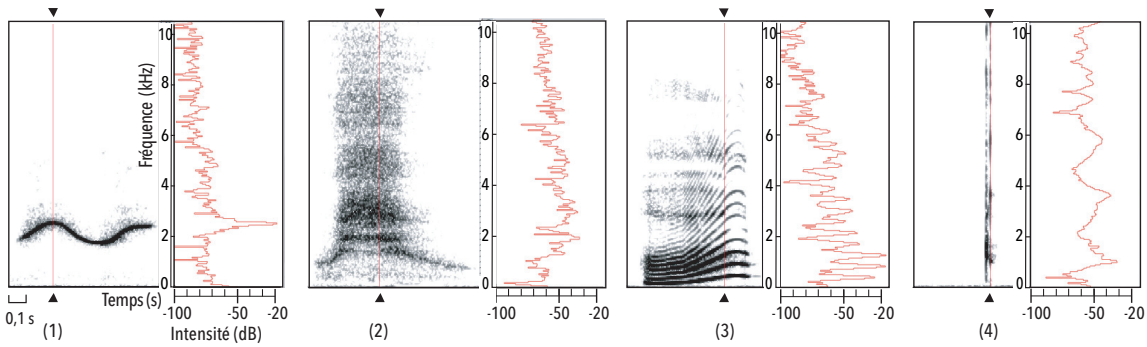


Figure 2.17 Les quatre sons de la figure 2.3. Le spectre I (dB)/*f* (kHz) est calculé à l'endroit indiqué par le curseur rouge sur le sonagramme. (Analyse Audiosculpt, fréquence d'échantillonnage = 44,1 kHz, 1024 pts, Δt = 23 ms)

La courbe rouge verticale s'affiche sur la même échelle de fréquence que le sonagramme. L'intensité en dB se lit par projection sur l'axe horizontal (échelle relative de 80 dB). De gauche à droite, on retrouve : (1) le sifflet dont le spectre montre une composante principale à environ 40 dB au-dessus du bruit de fond ; (2) le « chi » qui est un spectre continu mais sans énergie dans les basses fréquences ; (3) le son vocal qui comporte une dizaine d'harmoniques bien individualisés, les trois premiers étant les plus intenses ; (4) le premier claquement de langue, dont le spectre a été pris dans le rebond de l'explosion, est continu, avec trois colorations en fréquence, vers 1000, 3500 et 5500 Hz.

8. Pour avoir une démonstration pratique de l'incidence de la largeur de la fenêtre d'analyse sur le spectre, se reporter à la commande *Spectrum section* de Wavesurfer et modifier en direct le nombre d'échantillons ou « points » de la FFT.

Certains instruments (basson, hautbois) et, plus particulièrement la voix, ont des zones de résonance très marquées appelées « formants ». Il est intéressant de ne retenir que la courbe globale (ou enveloppe), en ignorant le détail des composantes isolées. La figure 2.18 montre un exemple avec indication des formants. (Pour la notion de formants, voir chapitre 9, § 1.4.2, page 453.)

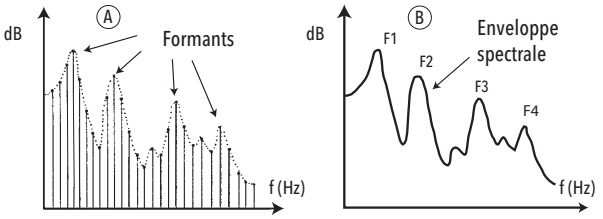


Figure 2.18 Analyse d'un spectre harmonique avec formants. En B, la courbe qui relie les maxima du spectre, l'enveloppe spectrale, met en évidence les formants.

3.5.2. Les spectres moyennés à long terme ou LTAS *(Long-Term Average Spectrum)*

Son 2.7 (13")

Gammes chromatiques, flûte, hautbois

Un spectre moyenné sur de longues durées temporelles fournit des informations sur la répartition statistique de l'énergie, indépendamment des hauteurs locales. Il correspond à l'appréciation qualitative des sons musicaux pour lesquels une écoute de longue durée est nécessaire. En effet, pour évaluer la sonorité d'un instrument, il faut parcourir toute la tessiture, jouer les sons graves et aigus et se faire une représentation mentale de l'équilibre général entre les différentes parties du spectre. Voici par exemple (figure 2.19), l'analyse comparée de deux instruments ayant une étendue similaire (la flûte traversière et le hautbois) jouant la même séquence, à la même intensité (Son **2.7**). Notez que l'échelle de fréquence du spectre est logarithmique ; la plus petite division est le tiers d'octave (soit une tierce majeure).

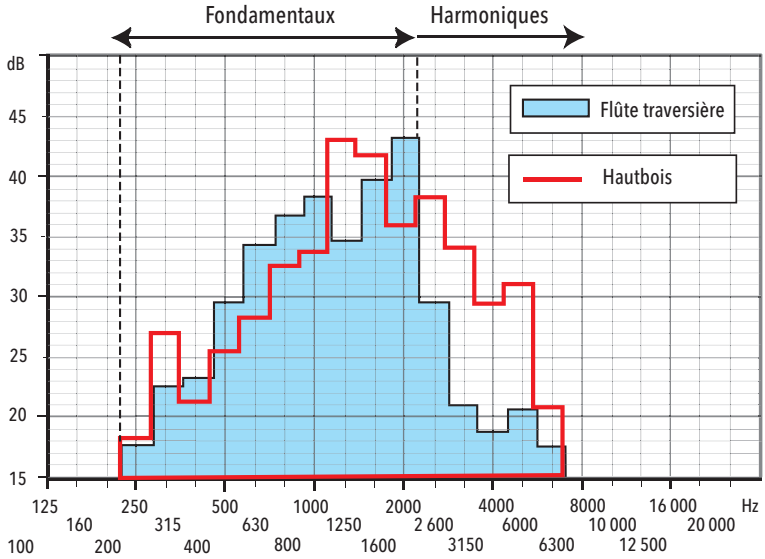


Figure 2.19 Spectres moyennés sur toute la durée d'une gamme chromatique ascendante, jouée sur une flûte traversière et sur un hautbois. Analyseur de spectre en tiers d'octave, Bruël & Kjaer. L'échelle verticale est en dB. L'échelle des fréquences est logarithmique.

On lit sur cette figure que le spectre moyenné du hautbois favorise les fréquences élevées. Bien que la flûte ait une étendue qui dépasse celle du hautbois, on note une différence significative de 10 dB dans la zone de fréquence des harmoniques compris entre 2600 et 6000 Hz. À l'inverse, la flûte dépasse le hautbois d'environ 5 dB dans les fréquences qui correspondent aux notes fondamentales.

Une autre méthode consiste à calculer la contribution énergétique de chaque bande spectrale par rapport à l'énergie totale, ce qui renseigne sur l'équilibre graves/aigus, indépendamment de la durée et indépendamment de l'intensité absolue. La figure 2.20 représente les spectres moyennés d'un Grand Fond d'orgue (Son **2.8a**), registration riche en fréquences graves, et du jeu de Fourniture (Son **2.8b**), registration riche en harmoniques aigus.

Orgue d'Ebersmunster. comparaison après restauration et avant restauration

Son 2.8a (48")

Grand Fond d'orgue

Son 2.8b (15")

Gamme diatonique sur la Fourniture du positif *(voir page 76)*

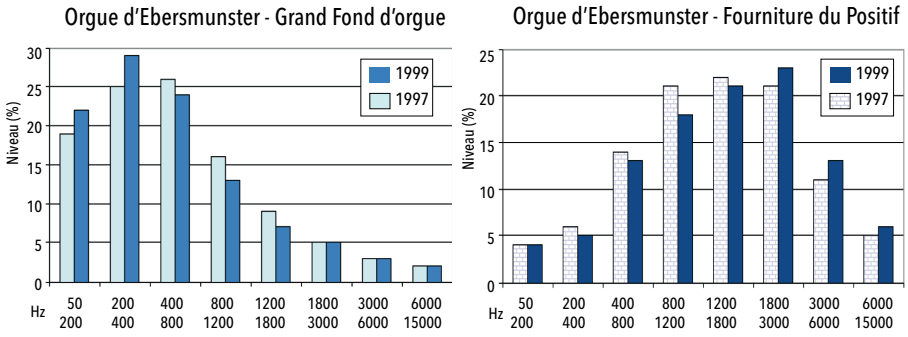


Figure 2.20 Spectre à long terme : comparaison de la sonorité de jeux d'orgue avant (zones claires) et après (zones sombres) restauration. Analyse IDS. Les chiffres de l'échelle horizontale indiquent les limites (en Hz) des bandes de fréquence dans lesquelles est effectué le calcul de l'énergie. L'échelle verticale est en pourcentage (%) de l'énergie totale.

Sur chaque figure on peut comparer les modifications, à deux années de distance, survenues après la restauration de l'orgue d'Ebersmunster. Le Grand Fond d'orgue a retrouvé sa plénitude dans les basses fréquences (Son **2.8a**) ; la Fourniture a gagné en brillance dans les très hautes fréquences (Son **2.8b**). L'analyse avec l'IDS développé par E. Leipp[9] divise l'étendue des fréquences en 8 bandes définies pour leur pertinence perceptive. La contribution de chaque bande est exprimée en pourcentage (%) de l'énergie totale.

3.6. Le traitement du signal et la resynthèse

Les analyses que nous avons présentées constituent une première approche et il est déjà possible d'en extraire un grand nombre de renseignements. L'étape suivante relève d'une discipline à part entière, le traitement du signal, que nous n'aborderons pas ici. La numérisation des sons autorise des traitements mathématiques complexes des données calculées par la transformée de Fourier ou par d'autres algorithmes d'analyse spectrale. Du simple point de vue de l'analyse, en effectuant des corrélations entre spectres successifs, on peut effectuer un suivi des composantes fréquentielles ou des formants. Il est possible aussi de pondérer les échelles de fréquence et d'intensité pour simuler, par exemple, le traitement de l'oreille interne sous forme de *cochléogrammes* (analyses proposées par le logiciel Praat).

9. IDS pour Intégrateur de densité spectrale. Voir Leipp, E., 1977b, *Bulletin du GAM* n° 94.

Selon nous, l'une des avancées les plus intéressantes du traitement du signal est le champ ouvert par les méthodes dites d'analyse-synthèse. L'analyse par FFT à court terme est réversible. Le calcul décompose le son en briques élémentaires auxquelles on peut appliquer des transformations fines, fréquentielles ou temporelles, qui seront prises en compte au moment de la resynthèse. Depuis peu, ces traitements, calculables en temps réel, invitent à explorer auditivement les sons. À peine une hypothèse est-elle formulée qu'il est aisé de la mettre en œuvre en agissant sur la représentation graphique et d'écouter le résultat dans les secondes qui suivent. De telles possibilités ouvrent de nouvelles perspectives de recherche pour comprendre les rapports entre l'écoute des sons et leur représentation.

Son 2.9 (24")

Expériences sur le son de bouche du transitoire (voir explications page 76)

L'exemple du Son **2.9** a été réalisé avec le logiciel Audiosculpt pour l'étude du transitoire d'attaque d'un tuyau d'orgue. L'analyse spectrale du son de ce tuyau (figure 2.21) montre que des sons inharmoniques – les sons de bouche – se produisent avant l'établissement du son fondamental du tuyau (A), qui peuvent jouer un rôle dans la qualité d'attaque perçue de ce tuyau (voir Castellengo, M., 1999).

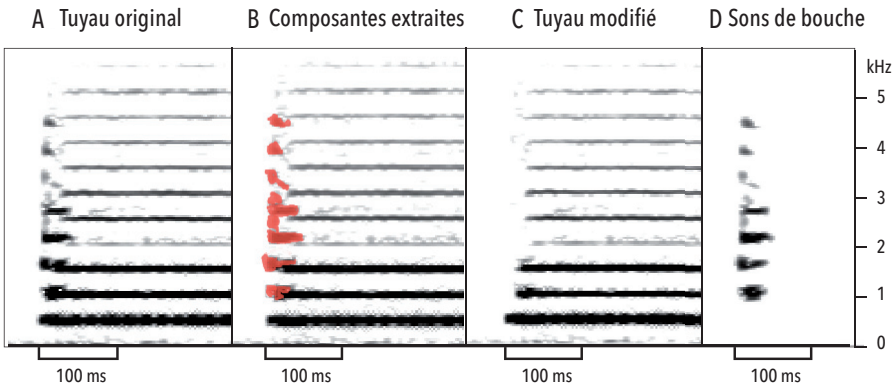


Figure 2.21 Analyse et synthèse par Phase Vocoder. A : son original du tuyau d'orgue. B : « gommage » manuel, sur l'image du sonagramme, des sons de bouche du transitoire. C : analyse spectrale du son du tuyau modifié. D : analyse des sons de bouche extraits de A.

Ce sont des phénomènes tellement brefs et complexes qu'il est difficile de bien les repérer à l'oreille. Pour mettre en évidence leur incidence perceptive supprimons ces sons de bouche et effectuons la resynthèse (B). Le tuyau démarre maintenant directement sur ses harmoniques (C). L'attaque du son paraît plus « molle », voire un peu terne. Écoutons isolément les sons de bouche qui ont été supprimés (D) : ils sonnent environ un demi-ton plus haut que le fondamental du tuyau (Son **2.9**), car ils correspondent à l'accrochage de modes propres qui sont plus hauts que les harmoniques du son stable (voir page 64). La comparaison entre le son original et le son transformé permet d'apprécier l'importance des sons de bouche dans la netteté et la clarté de l'attaque d'un tuyau d'orgue à bouche.

3.7. La représentation des intervalles musicaux

Une des difficultés rencontrées en analyse acoustique est celle d'une bonne correspondance entre l'échelle d'affichage des fréquences et la perception musicale. Deux possibilités existent : l'échelle linéaire et l'échelle logarithmique. La première correspond à la perception des intervalles musicaux, la seconde à la perception de

l'harmonicité des sons et permet en outre d'estimer les intervalles entre sons successifs par la méthode des harmoniques communs.

3.7.1. Échelles linéaires ou logarithmiques ?

Le système de notation musicale développé en Occident présente des similitudes avec le plan sonagraphique : la hauteur musicale est indiquée sur l'axe vertical de la portée, le temps (les rythmes) selon l'axe horizontal et l'intensité est spécifiée par des signes particuliers : indications de nuances (*p, f*) et de variations d'intensité (*cresc.*). Le tracé sonagraphique d'une mélodie en son sinusoïdal (Son **2.10**) se présente figure 2.22 comme une écriture analogue à celle de la musique, avec une plus grande précision pour les durées (tenues et silences) ; cependant sur cette figure les intervalles musicaux ne sont pas correctement représentés.

Son 2.10 (7")

Mélodie en son pur

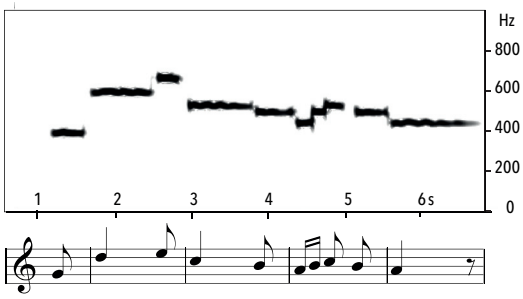


Figure 2.22 Analyse sonagraphique d'une mélodie et sa notation musicale. Le son pur a été obtenu par filtrage de l'harmonique 1 de la flûte traversière.

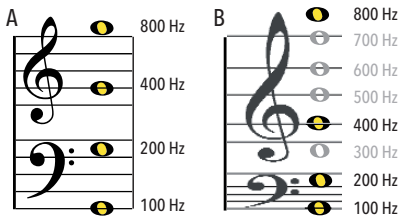


Figure 2.23 Comparaisons de deux échelles de représentation de la hauteur. A : celle des intervalles sur la portée musicale qui correspond à une échelle logarithmique des fréquences. B : celle des harmoniques, qui est linéaire en fréquence, mais correspond à une distorsion visuelle de la portée musicale, car les intervalles sont comprimés dans les fréquences basses et agrandis dans l'aigu.

Écrivons les octaves du sol_1 sur une portée musicale et plaçons en vis-à-vis leurs fréquences approchées : 100, 200, 400, 800 Hz. On voit (figure 2.23-A) que les fréquences associées aux notes de la portée positionnées verticalement de façon équidistante, sont dans un rapport constant : en l'occurrence celui de l'octave qui est 2. Il s'agit d'une échelle « logarithmique » en fréquence (voir Glossaire, page 521). À l'opposé, l'échelle des fréquences d'un sonagramme est habituellement « linéaire », car l'analyse FFT est effectuée à Δf constant : l'écart entre graduations successives contient le même nombre de hertz. Pour établir la correspondance entre la notation musicale et une échelle de fréquences linéaire, il faut anamorphoser la portée musicale (figure 2.23-B). On constate que l'intervalle d'octave grandit du grave à l'aigu, mais que, cette fois, les notes correspondant aux harmoniques sont équidistantes.

Son 2.11 (7″)

La mélodie jouée à la flûte traversière

Considérons maintenant l'analyse au sonagramme de la mélodie jouée à la flûte traversière (Son **2.11**) avec tous ses harmoniques.

Avec l'échelle linéaire des fréquences (figure 2.24 gauche), les harmoniques de la flûte sont immédiatement repérables par l'équidistance des traits superposés. Lorsqu'on opte pour un affichage en échelle logarithmique, ce qui nécessite une transformation optique de l'image (figure 2.24 droite), on remarque que la mélodie du fondamental est facilement lisible.

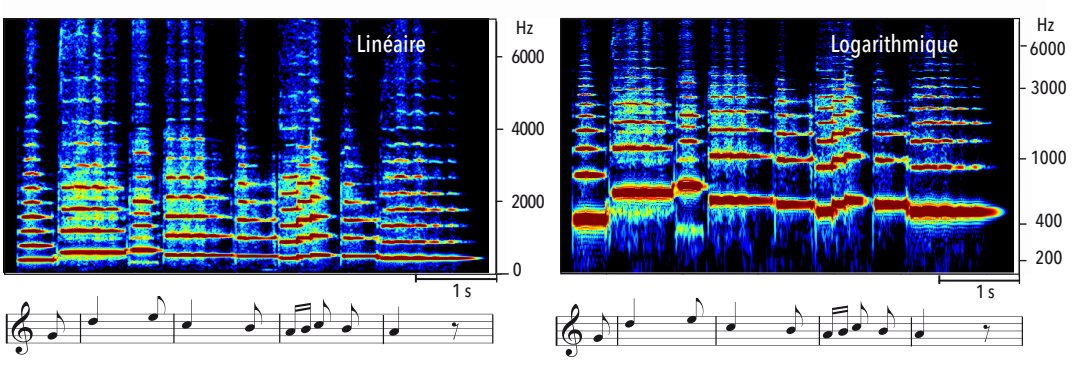


Figure 2.24 Analyses sonagraphiques comparées d'une mélodie jouée à la flûte traversière en échelle linéaire des fréquences (gauche) et en échelle logarithmique des fréquences (droite).

En échelle linéaire la ligne mélodique du fondamental est « écrasée » dans le bas de la figure, mais l'équidistance des harmoniques est aisément repérable. Si l'on choisit un affichage logarithmique des fréquences, la zone des fondamentaux graves est agrandie, mais l'équidistances des raies disparaît. Selon ce que l'on souhaite mettre en évidence, il faudra donc choisir entre une représentation musicale des intervalles ou une représentation physique des données spectrales.

3.7.2. Les intervalles musicaux et les coïncidences d'harmoniques

Une ressource précieuse pour apprécier les intervalles entre deux sons successifs riches en harmoniques est la recherche des harmoniques communs.

Deux sons à l'octave ont un harmonique sur deux à la même fréquence : H2 du son grave coïncide avec H1 du son aigu. Pour un intervalle de quinte ascendante, H3 du son grave coïncide avec H2 du son aigu, et ainsi de suite (voir chapitre 1, § 4.4). La représentation linéaire des fréquences permet de repérer la coïncidence en fréquence même lorsque des harmoniques manquent, et en particulier lorsque le fondamental est filtré, puisque c'est l'équidistance des raies spectrales qui est caractéristique de la hauteur du son.

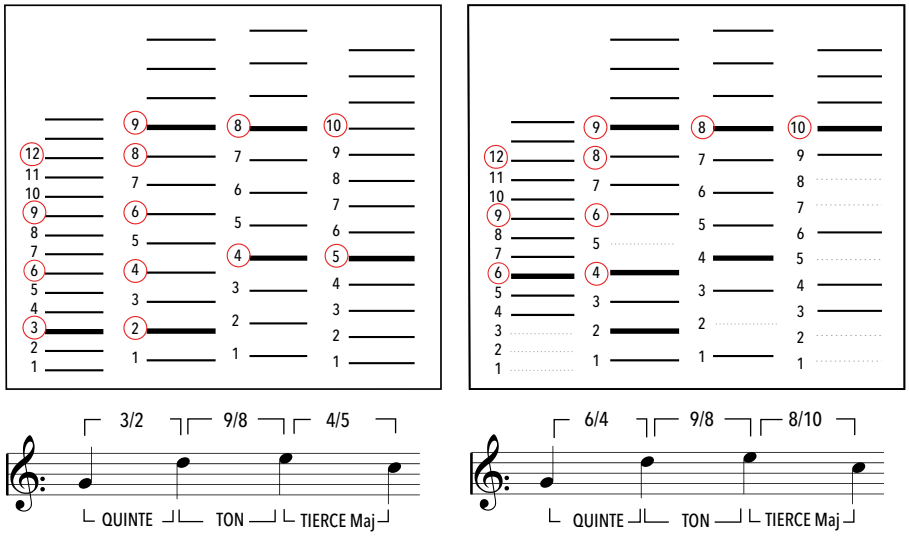


Figure 2.25 Schéma des harmoniques de quatre notes représentées selon une échelle linéaire des fréquences. Étant donné que les harmoniques d'un son sont équidistants et que les harmoniques communs à deux sons consécutifs ont la même fréquence (traits renforcés), on peut apprécier l'intervalle entre deux sons, même lorsque le spectre est incomplet : voir sur la figure de droite le filtrage des premiers harmoniques du 1er et du 4e son.

3.7.3. L'agrandissement des intervalles dans l'aigu (échelle linéaire)

Une autre caractéristique à prendre en compte lors de l'affichage linéaire des fréquences est le fait que les mouvements mélodiques s'agrandissent sur les harmoniques de rang élevé (Son **2.12**). À première vue, l'effet est troublant. En observant le mouvement mélodique des deux harmoniques surlignés en rouge sur la figure 2.26, on constate que le vibrato est deux fois plus ample sur l'harmonique 8 que sur l'harmonique 4. Effectivement, lors d'une variation mélodique, les écarts en hertz sont multipliés par le rang de l'harmonique. Cependant, le rapport de fréquence des composantes de même rang étant caractéristique d'un intervalle donné, on comprend que celui-ci reste constant pour l'oreille.

Son 2.12 (8")
Chant (soprano)

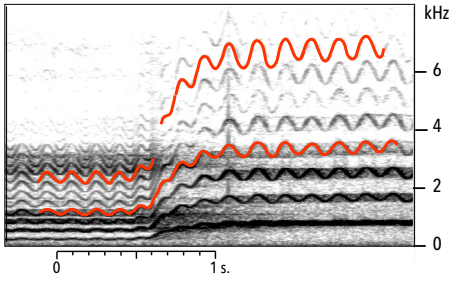


Figure 2.26 Extrait de l'analyse d'une mélodie chantée mettant en évidence l'agrandissement des dessins mélodiques avec le rang de l'harmonique dans la représentation linéaire des fréquences.

Le zoom visuel du tracé est intéressant lorsque les mélodies vocales ou instrumentales ont peu d'énergie dans les basses fréquences, ce qui est fréquent dans les

musiques traditionnelles. On peut alors suivre la ligne mélodique sur un harmonique à l'octave du fondamental (H2 ou H4), ce qui fournit aussi plus de précision.

3.7.4. L'interprétation des mesures

Les logiciels d'analyse du son procurent des mesures d'intensité et de fréquence, encore faut-il en évaluer la pertinence. Dans un spectre dB/Hz, les valeurs d'intensité sont relatives. Le plus souvent, la composante de plus grande amplitude a par convention la valeur de 0 dB et les autres s'affichent en valeurs négatives. L'analyse fournit donc les rapports d'intensité entre les différentes composantes du spectre.

La mesure de la fréquence est beaucoup plus délicate à interpréter. Il ne faut jamais abandonner l'estimation à l'oreille, en s'aidant d'un diapason, ce qui permet souvent de corriger les erreurs d'octave que l'on trouve dans quelques logiciels. Lorsqu'il s'agit d'une analyse par la transformée de Fourier, il faut garder à l'esprit que la valeur affichée sur un spectre doit être pondérée par la largeur de la fenêtre d'analyse. Si, par exemple, l'analyse est faite avec 2048 échantillons à une fréquence d'échantillonnage de 44,1 kHz, la fréquence mesurée est valable à ± 22 Hz. Comme cette incertitude est constante, l'erreur d'estimation est d'autant plus grande que la fréquence à mesurer est faible, tout particulièrement les sons de basse fréquence.

On retiendra comme règle pratique qu'une mesure de la fréquence avec quatre chiffres significatifs est suffisante en acoustique musicale pour rendre compte de la perception. Les chiffres après la virgule n'ont de sens que pour les sons de fréquence inférieure à 1000 Hz. Ainsi 1577 Hz, 440,3 Hz, 85,42 Hz sont des estimations équivalentes en précision pour rendre compte d'une fréquence fondamentale.

4. Quelques exemples d'analyses

Dans les pages qui suivent, nous proposons, en forme de synthèse, quelques exemples pratiques d'analyses appliquées à une sélection de sons avec pour objectif :

- les diverses représentations d'un même phénomène sonore (figures 2.27, 2.29 et 2.30),
- quelques expériences simples portant sur les relations entre partiels et harmoniques d'un instrument de musique (figures 2.32 à 2.35),
- l'incidence des paramètres de l'analyse sur la représentation visuelle (figures 2.36 et 2.37),
- la comparaison d'une même phrase musicale chantée et jouée au piano (figure 2.38),
- une introduction au déchiffrage de scènes sonores complexes (figures 2.39 et 2.40).

Le chapitre se termine avec l'explication détaillée des analyses comparées d'un son de hautbois réalisées avec trois logiciels d'analyse spectrale (Wavesurfer, Praat et Audiosculpt).

4.1. Un cas d'école : sonagrammes, formes d'ondes et spectres de sons électroniques

La figure 2.27 montre l'analyse de trois signaux classiques en synthèse sonore: l'*onde sinusoïdale*, qui correspond à une vibration dont les alternances sont graduelles, sans à-coup ; l'*onde carrée,* qui traduit une oscillation passant brusquement (parties verticales de l'onde) d'un état à un autre (parties horizontales de

l'onde) ; l'onde *en dents de scie*, qui correspond à un mouvement graduel dans un sens (oblique), suivi d'un retour quasi instantané (partie raide). On remarquera que la forme d'onde s'écarte du schéma théorique du générateur, représenté en traits pointillés. Le signal enregistré a subi, tout au long de la chaîne d'amplification, différentes distorsions de phase qui modifient la forme de l'enveloppe temporelle, sans affecter son contenu spectral (voir *Phase*, page 527 du glossaire).

Chaque forme d'onde périodique peut être décomposée en ses harmoniques. L'onde sinusoïdale ne comporte qu'une composante ; l'onde carrée est constituée des seuls harmoniques impairs ; l'onde en dents de scie comprend toute la série harmonique, avec des amplitudes qui décroissent avec le rang. Ces résultats se démontrent aisément en utilisant la technique pratique d'addition des harmoniques.

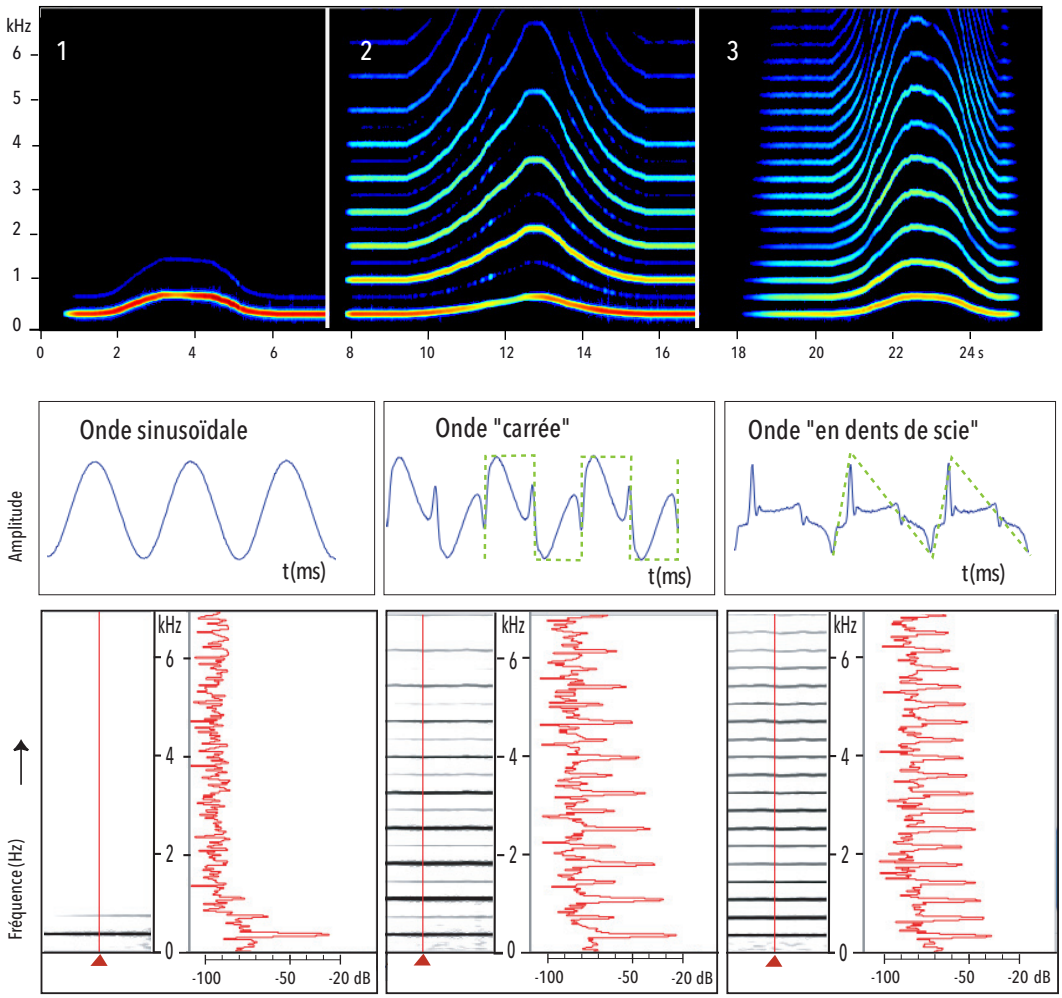


Figure 2.27 Trois formes d'ondes classiques en électronique. Partie supérieure : sonagramme d'un glissando ascendant/descendant. Partie médiane : forme d'onde du signal sonore enregistré et analysé. Partie inférieure : sonagramme de la partie finale de chaque glissando et spectre amplitude/fréquence calculé au curseur rouge.

Son 2.13a (26″)

Glissandos 1, 2 et 3

Son 2.13b (4″)

Note tenue terminale de chaque glissando

Ces trois sortes d'ondes, reconnaissables à l'oreille – écoutez les sons **2.13a** et **2.13b** – peuvent être mises en relation avec le mouvement des excitateurs de certains instruments de musique. Ainsi, le son pur de l'onde sinusoïdale paraît flûté, comme celui qui est produit par un jet d'air à l'entrée d'un résonateur (par exemple, une bouteille). Le son de l'onde carrée évoque la sonorité du registre grave de la clarinette, lorsque le mouvement de l'anche passe brusquement de la position ouverte à la position fermée. L'onde triangulaire évoque un son de corde frottée. Dans son mouvement vibratoire, cette dernière est d'abord entraînée lentement par l'archet, revient brusquement, puis est de nouveau entraînée par adhérence. L'évolution temporelle du déplacement local de la corde sous l'archet est ainsi fortement dissymétrique et la courbe qu'elle décrit est dite « en dents de scie ».

La mémorisation du caractère de ces trois sonorités est un bon repère pour apprendre à analyser à l'oreille le contenu harmonique de la plupart des sons.

4.2. L'anatomie d'un son de steel-drum

Son 2.14 (5″)

Une note de steel-drum

Nous avons rassemblé, figures 2.29 et 2.30 (p. 60 et 61), diverses analyses d'un son de steel-drum (note $ré_2$, Son **2.14**).

Le steel-drum (aussi appelé steel-pan, figure 2.28), instrument provenant des îles Caraïbes, possède une sonorité très particulière du fait des nombreuses interactions se produisant entre les différentes zones de la surface du *pan*.

Figure 2.28 Un steel-drum vu du dessus, et note du Son 2.14

D'après Monteil, M., 2010, DEA, figure 1.

Cet instrument appartient à la catégorie des percussions, ce que l'on reconnaît immédiatement sur le sonagramme (3) par la présence du trait vertical indiquant la frappe, suivi de fréquences non harmoniques et d'évolution complexe. Cependant, à l'écoute, le transitoire d'attaque n'est pas immédiatement suivi d'une décroissance comme c'est habituellement le cas dans les instruments à percussion. Il arrive même souvent qu'une « remontée » sensible de l'amplitude sonore se fasse sentir dans les cent premières millisecondes, donnant parfois l'impression d'un son quasi entretenu.

Sur les deux premières analyses de la figure 2.29 – la courbe d'intensité en dB (1) et le signal temporel (2) –, on peut voir que la montée rapide de l'intensité au moment de l'attaque se poursuit pendant les 300 premières millisecondes. Cette caractéristique est très visible sur la représentation tridimensionnelle (6), où la mise en perspective de la variation d'amplitude des quatre premiers partiels pendant la première seconde de son (5) rend sensible la complexité du transitoire d'attaque.

Le graphique (5) permet de préciser le déroulement temporel : le partiel 1 s'établit en premier, suivi du partiel 3, puis du 2 et enfin du 4. Comme tous les sons produits par percussion, le spectre d'un son de steel-drum est en constante évolution. Nous avons réalisé deux spectres instantanés (4) aux instants A et B du sonagramme, séparés seulement par 700 ms. En B, on constate que les partiels aigus, au-delà de 600 Hz, chutent plus rapidement que les fréquences graves.

À l'oreille, la note de ce son est un *ré*$_2$ (146 Hz). On voit effectivement une pseudo périodicité dès le début du signal temporel (7) pendant la montée isolée du partiel 1. Une mesure grossière de la durée des dix premières oscillations fournit une estimation de la période de l'ordre de 6,5 ms. En ciblant une mesure de la fréquence fondamentale entre 100 et 220 Hz, le logiciel Praat affiche 147,7 Hz, ce qui est proche de la fréquence d'un *ré*$_2$. Toutefois la sensation de hauteur d'un son inharmonique comme l'est le steel-drum ne peut, en toute rigueur, être assimilée à la fréquence de la composante la plus grave. Nous renvoyons le lecteur au chapitre 6, § 3.4, où plusieurs exemples de sons inharmoniques sont analysés (piano, cloche-tube, timbale, etc.)

Pour satisfaire la curiosité de certains lecteurs, nous avons fait la mesure des premiers partiels après les avoir filtrés individuellement. Le rapport Pn/P1 permet d'estimer l'écart des partiels par rapport à une série harmonique (nombres entiers).

On remarque que les partiels 2 et 6 sont doubles[10]. Dans l'ensemble, les fréquences propres des partiels sont assez voisines de celles des harmoniques du partiel 1, à l'exception du partiel 5 qui est beaucoup plus bas que l'harmonique 5 de P1. L'instrument est donc bien accordé.

Tableau 1. Mesure de la fréquence des partiels du steel-drum

N° partiel	P1	P2	P3	P4	P5	P6	P7
Fréquence	147,7	247-291	436,3	589,2	678,5	883-875	1024
Pn/P1	**1**	**2,01 et 1,97**	**2,95**	**3,98**	**4,59**	**5,97 et 5,92**	**6,93**

10. Parmi les explications, on peut invoquer le couplage avec les zones des notes voisines : *ré*$_3$ et *la*$_4$.

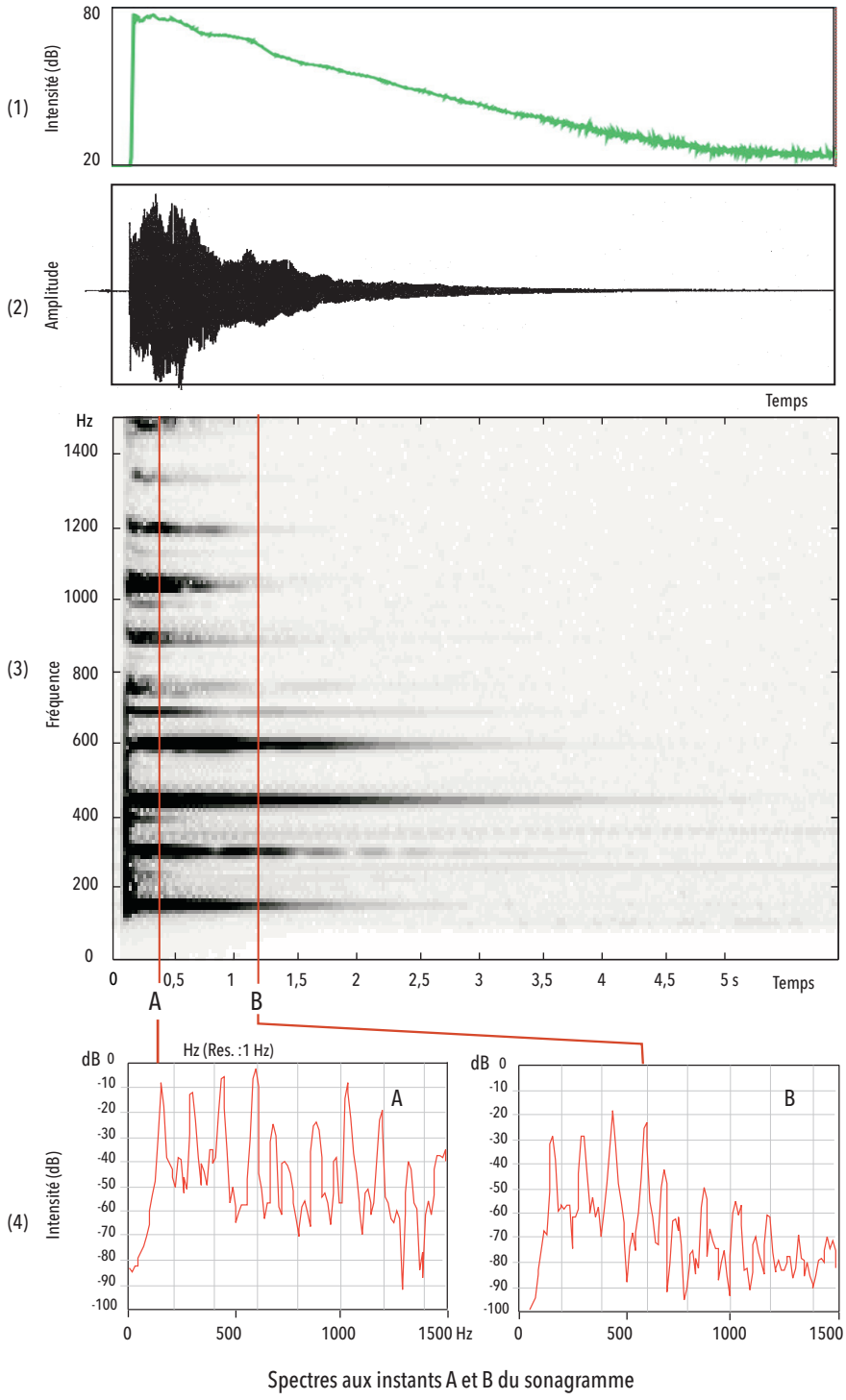
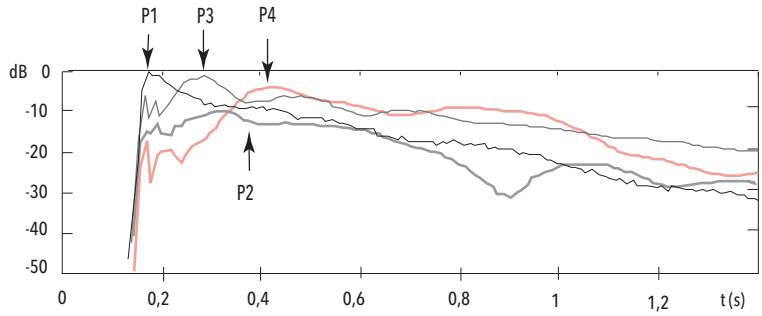


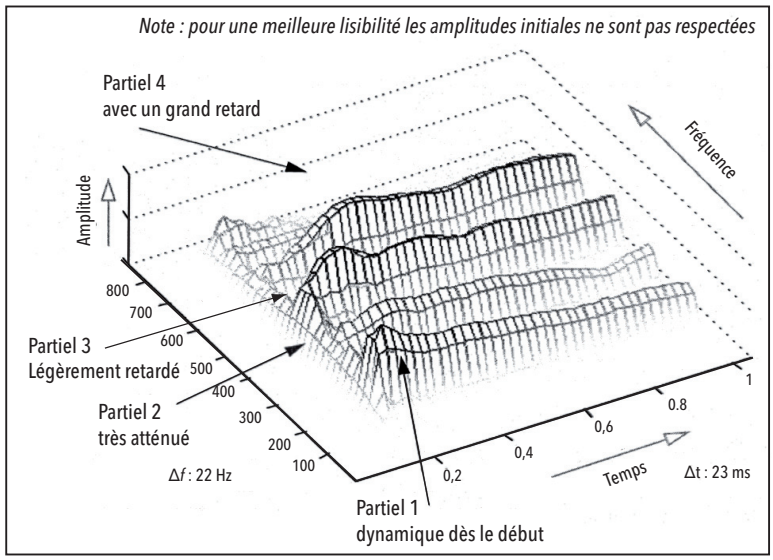
Figure 2.29 Analyse d'un son de steel-drum ($ré_2$ = 147,7 Hz). (1) Intensité globale. (2) Signal temporel. (3) Sonagramme : FFT de 4 096 points (92,8 ms). (4) Spectres aux instants A et B.

D'après Gaillard, P., 2000, Thèse, pages 28 à 32.

(5)



(6)



(7)

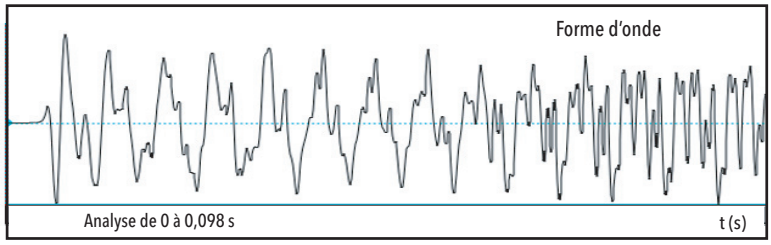


Figure 2.30 Analyse d'un son de steel-drum ($ré_2$ = 147,7 Hz). (5) Évolution temporelle de l'intensité des quatre premiers partiels. (6) Spectrogramme tridimensionnel (fréquence, intensité, temps). (7) Forme d'onde des 100 premières millisecondes.

D'après Gaillard, P., 2000, Thèse, pages 28 à 32.

Son 2.15 (27″)

Verre (voir p. 77)

4.3. Le verre percuté et l'entretien des deux premiers modes propres

Lorsqu'on donne une légère impulsion au verre, il « sonne » (Son **2.15**). Du point de vue physique, le verre se déforme selon ses modes vibratoires potentiels, chacun émettant sa fréquence propre (voir Glossaire, *Mode vibratoire (mode propre)*).

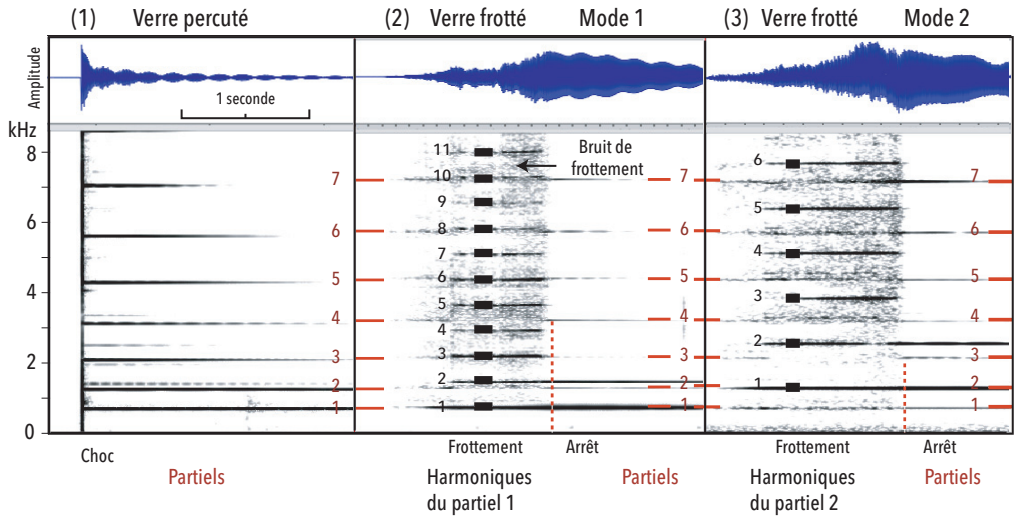


Figure 2.31 Analyses comparées d'un même corps excité de façon impulsionnelle : verre percuté, puis entretenu avec un archet de violon, successivement sur le partiel 1 (fa_4 = 725 Hz) et le partiel 2 (mib_5 = 1287 Hz). La position fréquentielle des partiels et leur numérotation est indiquée en rouge. Les numéros des harmoniques du son entretenu sont en noir. Remarquez les bruits de frottement apparaissant dans l'aigu du spectre. (Analyse Audiosculpt, 2 048 pts, soit Δf = 21,5 Hz)

L'analyse acoustique du verre percuté (figure 2.31-1) est caractéristique :

- une attaque brève avec montée rapide de l'amplitude (courbe bleue) et trait vertical sur l'analyse spectrale du sonagramme ;
- immédiatement après l'attaque, émission de fréquences pures sans relation harmonique entre elles : ce sont les partiels numérotés de 1 à 7 sur l'analyse. Certains sont instables et produisent des battements rapides apparaissant sous forme de traits pointillés (partiels 2 et 4).

Avec un archet, excitons le verre sur sa fréquence la plus grave (figure 2.31-2), celle du mode 1. La vibration s'établit lentement, avec une sonorité un peu grinçante. De fait, pendant les premières millisecondes, les différents partiels en compétition tentent de s'imposer : on voit leur trace aux mêmes fréquences que celle du verre percuté. Puis, brusquement, le partiel 1 l'emporte et, pendant un bref instant (de 0,7 à 1,3 s), de nombreux harmoniques (11) sont visibles, numérotés en noir sur la figure. L'entretien, accompagné d'un important bruit de frottement, est difficile à stabiliser ; l'archet quitte le bord du verre. L'extinction du son montre une amplitude importante du partiel 1 et la réapparition des autres partiels éliminés par l'entretien.

L'excitation du mode 2, plus difficile à accrocher, s'établit néanmoins, mais avec une persistance visible des fréquences des modes propres, ce qui accuse encore plus le caractère grinçant du son (figure 2.31-3).

Les fréquences des modes propres du verre émises lors de l'excitation par choc, mesurées avec Praat après filtrage de chaque composante, sont portées dans le tableau ci-dessous. L'écart entre ces fréquences (1re ligne) et celles des harmoniques du mode 1 entretenu (2e ligne) est particulièrement important, ce qui explique les difficultés rencontrées pour entretenir la vibration du verre avec un archet.

Tableau 2. Fréquence des composantes spectrales du verre B, « fa_4 + 65 cents » (ou $fa\#_4$ – 35 cents)

Numéros des composantes	1	2	3	4	5	6	7
Excitation par choc : fréquences (Hz) des modes propres du verre (partiels)	725,5	1281 1279	2137	3182	4372	5702	7166
Excitation entretenue : fréquences (Hz) des harmoniques du mode 1	722 *(722)*	1444 *(722 × 2)*	2166 *(722 × 3)*	2888 *(722 × 4)*	3610 *(722 × 5)*	4332 *(722 × 6)*	5054 *(722 × 7)*

Cette expérience illustre de façon manifeste les interactions entre les fréquences des modes propres et l'établissement du son entretenu. Dans l'univers des musiques occidentales, le savoir-faire des luthiers a conduit à transformer peu à peu la géométrie des instruments et les propriétés des cordes, de sorte que les modes propres soient au mieux alignés sur les harmoniques du mode 1, qui est le plus grave. Lorsque ce n'est pas possible, comme par exemple pour l'émission de certaines notes aiguës de la clarinette ou du hautbois, les problèmes d'attaque sont souvent à imputer à l'inharmonicité des modes.

4.4. L'octaviation et le quintoiement

Une flûte octavie, une clarinette « quintoie » (Son **2.16**). Ces termes décrivent l'intervalle musical que l'on peut produire sur un instrument à vent quand on modifie l'excitation (souffle, pincement) sans changer le doigté. Le deuxième mode vibratoire ainsi obtenu est un nouveau fondamental pourvu d'harmoniques (voir chapitre 1, § 5.2.2). Du point de vue sonore, il présente une particularité remarquable très visible sur l'analyse de la figure 2.32 ci-après.

Son 2.16 (11")

Partiels des modes 1 et 2 (flûte, clarinette)

Dans l'exemple de la flûte traversière, l'écartement des raies du son à l'octave supérieure semble très similaire au premier son. Pourtant, la numérotation des harmoniques permet de remarquer que ceux-ci ont une intensité plus grande et que les raies intermédiaires sont pratiquement des traces de bruit, aux fréquences des harmoniques du mode 1. En réécoutant l'exemple sonore on remarque clairement la présence d'un son « chuchoté » plus grave, à l'octave inférieure, c'est-à-dire à la hauteur du mode 1 ou fondamental do_3 de l'instrument, présent pendant toute la durée du do_4.

Le deuxième mode vibratoire d'un tuyau cylindrique à anche comme la clarinette est à un intervalle de douzième, c'est-à-dire une octave + une quinte. Ce son riche en harmoniques présente également la trace de fréquences inférieures entre les harmoniques, principalement au moment du transitoire. Pour ajuster précisément

Son 2.16 (11")

Partiels des modes 1 et 2 (flûte, clarinette)

la fréquence des partiels, réduire les sons indésirables et faciliter l'attaque, les instrumentistes disposent de clés ouvrant de petits trous dits « trous de registre ». Un trou de registre est positionné à l'emplacement d'un nœud de pression du premier mode, ce qui a pour effet d'en gêner la production et donc d'éliminer les traces sonores que l'on entend dans notre exemple.

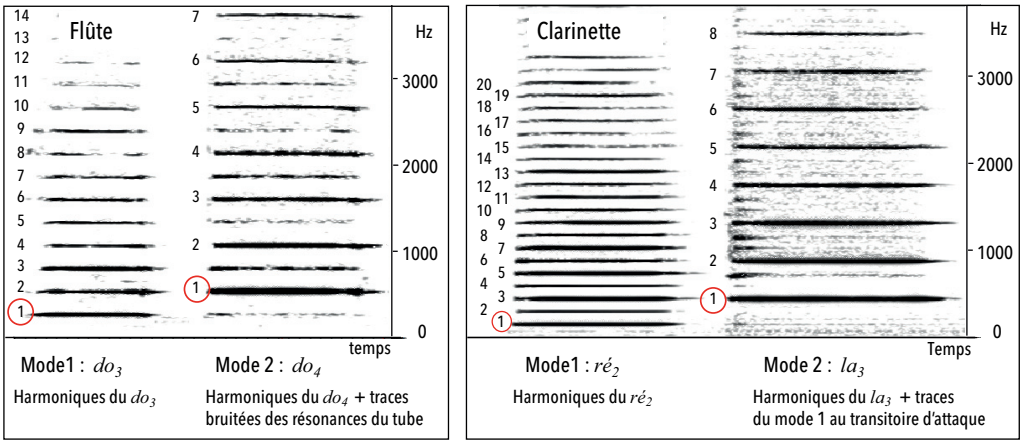


Figure 2.32 Analyse spectrale des sons produits par l'entretien des deux premiers modes vibratoires d'un instrument octaviant (la flûte traversière) et d'un instrument quintoyant (la clarinette). Dans ces deux exemples l'instrumentiste joue le mode 2 sans utiliser les trous de registre qui ont pour rôle d'éliminer les traces sonores du mode 1, visibles sur la figure.

4.5. La séquence complète des partiels d'une flûte et d'une clarinette

Son 2.17 (14")

Flûte traversière (doigté do$_3$)

Cet exemple reprend et complète celui du paragraphe précédent. Sur une flûte traversière dont on bouche tous les trous latéraux produisons – en modifiant la pression du souffle, la direction du jet et le recouvrement de l'embouchure – la suite des partiels P1, P2, P3... P8 qui correspondent aux modes propres du tuyau. Dans notre exemple, on a : do_3, do_4, sol_4, do_5, mi_5, etc., soit une suite analogue à celle des harmoniques du do_3. Toutefois, on remarque à l'écoute (Son **2.17**) que les intervalles s'agrandissent au fur et à mesure que l'on progresse dans l'aigu : les fréquences des partiels ne sont pas des multiples entiers de la fréquence du partiel 1 comme le montre l'exemple numérique :

Fréquences (Hz) des partiels produits sur le do_3 d'une flûte traversière mesurées avec Praat :

P1 = 263,5 ; P2 = 529,7 ; P3 = 794,5 ; P4 = 1 061 ; P5 = 1 324 ; P6 = 1 590 ; P7 = 1 861 ; P8 = 2 132

Le partiel 4 de la flûte (1 061 Hz) est plus haut que l'harmonique 4 du do_3 (H4 = 263,5 × 4 = 1 052 Hz)

La même expérience effectuée avec une clarinette en *sib* permet de jouer six partiels en modifiant le réglage de l'embouchure et la pression du souffle (Son **2.18**). P1 est un $ré_2$ (hauteur réelle), P2 un la_3, P3 un $fa\#_4$. Cette fois la suite des partiels avoisine la suite des harmoniques impairs du $ré_2$ et, au fur et à mesure que l'on progresse dans l'aigu, la hauteur des partiels s'abaisse.

Son 2.18 (19")

Clarinette (son fondamental réel, $ré_2$)

Fréquences (Hz) des partiels produits sur le $ré_2$ d'une clarinette en *sib* mesurées avec Praat :

P1 = 145,2 ; P2 = 433,2 ; P3 = 707,7 ; P4 = 945,4 ; P5 = 1 345 ; P6 = 1 593

Le partiel 2 de la clarinette (433,2 Hz) est plus bas que l'harmonique 3 du $ré_2$ (H3 = 145,2 x 3 = 435,6 Hz)

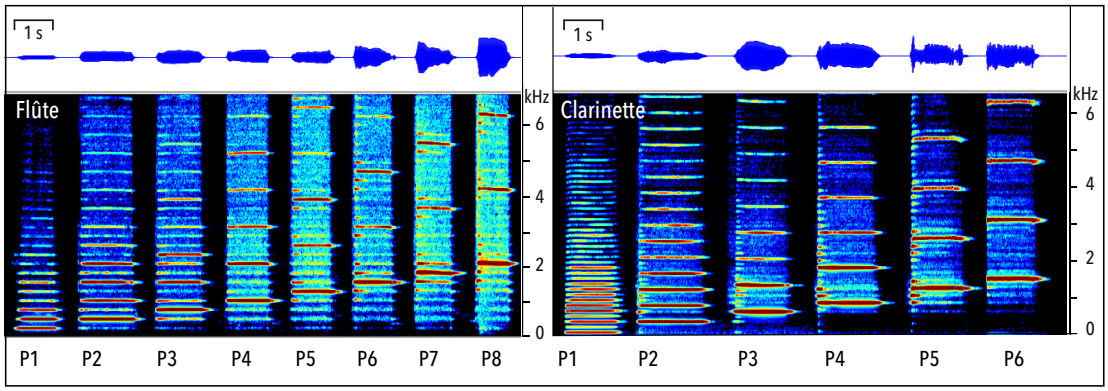


Figure 2.33 Partiels successifs de deux instruments à vent à trous, joués sur le doigté du son le plus grave (tous les trous bouchés). Le premier son, P1, est le son fondamental du tuyau. Chaque partiel est un son riche en harmoniques puisque l'oscillation est entretenue. Entre les harmoniques on remarque la persistance de raies intermédiaires correspondant aux harmoniques de P1 dont la hauteur subsiste sous forme de « son chuchoté » dans la flûte, et de « canard » discret à l'attaque du son de la clarinette.

4.6. Une mélodie jouée à la tilinca roumaine

Nous venons de voir que les modes vibratoires des tuyaux à embouchure de flûte de fine taille, comme la flûte traversière[11], sont quasi-harmoniques : la succession des sons que l'on peut jouer sans changer le doigté fournit la suite bien connue des intervalles : octave, quinte, quarte, etc. La tilinca roumaine, instrument sans trous latéraux, fonctionne sur ce principe (chapitre 1, § 6.2).

Si l'écoute en est claire (Son **2.19**), la lecture de l'analyse apparaît plus difficile (figure 2.34 ci-après), car le son est très bruité. Pour trouver la mélodie, il faut rechercher le rang du partiel joué. Le schéma de droite, placé en vis-à-vis de l'analyse des deux premières secondes, fournit un guide pour le dépouillement de la mélodie. L'instrumentiste dispose de deux séries de partiels : ceux du tuyau ouvert (O) en noir (fondamental 228 Hz, sib_2) et ceux du tuyau bouché (B) en rouge (fondamental 114 Hz, sib_1) pour les appoggiatures et les sons intermédiaires.

Son 2.19 (51")

Jeu de la flûte tilinca

On lit : brève appoggiature du 5e partiel du tuyau bouché (do_5) et note longue du 5e partiel du tuyau ouvert ($ré_5$), puis brève appoggiature double du 7e partiel bouché (sol_5) + 7e partiel ouvert (lab_5), suivie du 6e partiel ouvert (fa_5), et ainsi de suite.

11. La taille est le rapport du diamètre à la longueur du tube. Elle vaut (mesures en cm) 60/1,8 = 33 pour la flûte traversière, et 60/1,2 = 50 pour la tilinca.

Son 2.19 (51″)

Jeu de la flûte tilinca

Comme pour la flûte traversière, la série complète des résonances du tube est visible et audible sous forme de son chuchoté. Le bouchage de l'extrémité inférieure du tuyau s'accompagne d'un bruit grave de percussion, de durée brève, par lequel le musicien rythme son jeu (voir flèches sur la figure 2.34).

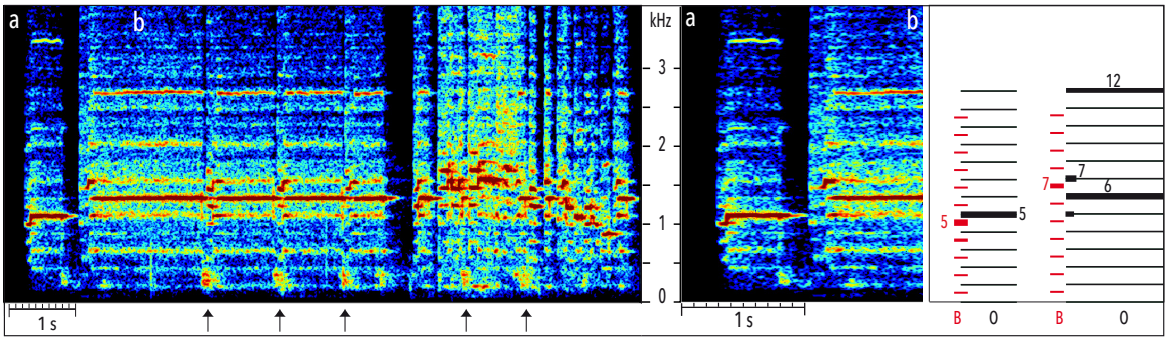


Figure 2.34 Spectrographie de la mélodie jouée à la tilinca. À gauche, le sonagramme du début (9″). À droite, l'analyse agrandie des deux premières secondes de jeu (section ab) et un schéma expliquant les partiels joués tantôt sur le tuyau bouché (B, rouge) tantôt sur le tuyau ouvert (O, noir). Les flèches indiquent le bruit de percussion du rebouchage bref pratiqué à l'attaque des sons (FFT : 3 000 pts).

4.7. Les mélodies d'harmoniques et les mélodies de partiels

La figure 2.35 montre l'analyse au sonagramme de deux exemples caractéristiques (Sons **2.20** et **2.21)** déjà entendus au premier chapitre (page 25 et page 22).

Son 2.20 (9″)

Tuba : partiels

Son 2.21 (7″)

Chant diphonique

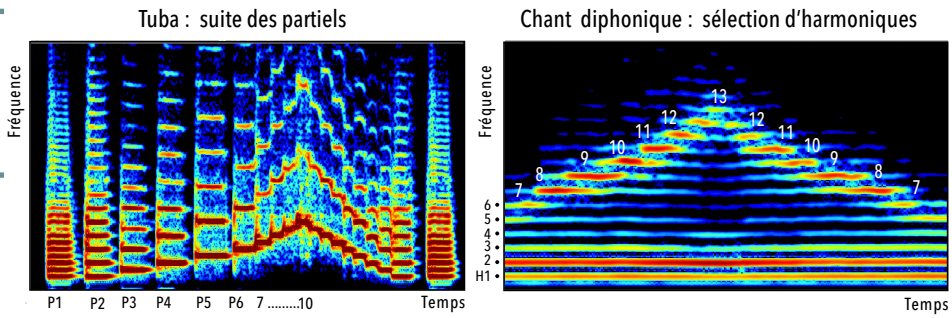


Figure 2.35 Mélodie ascendante et descendante. À gauche, succession des partiels d'un instrument à embouchure (tuba). À droite, sélection d'harmoniques d'un son vocal par résonance buccale.

L'analyse de la série des partiels du tuba (figure 2.35, à gauche), jouée en série ascendante et descendante montre que pour chacun des modes vibratoires de l'instrument joués successivement, on obtient un son périodique, riche en harmoniques.

L'analyse de la mélodie de l'exemple de chant diphonique (figure 2.35, à droite) fait apparaître la sélection ascendante et descendante des harmoniques du spectre que le chanteur opère en modifiant la fréquence de résonance de sa cavité buccale. Pendant toute la durée du son, la hauteur fondamentale du son vocal est constante, ce que montre l'horizontalité du spectre harmonique.

4.8. L'analyse de la voix humaine : formants des voyelles ou intonation ?

La voix parlée est une bonne illustration de l'importance du choix des paramètres d'une analyse, particulièrement celui de la résolution en fréquence. Soit la phrase « Bonjour monsieur », prononcée par un homme (Son **2.22**, figure 2.36).

Son 2.22 (2″)

Voix parlée (homme)

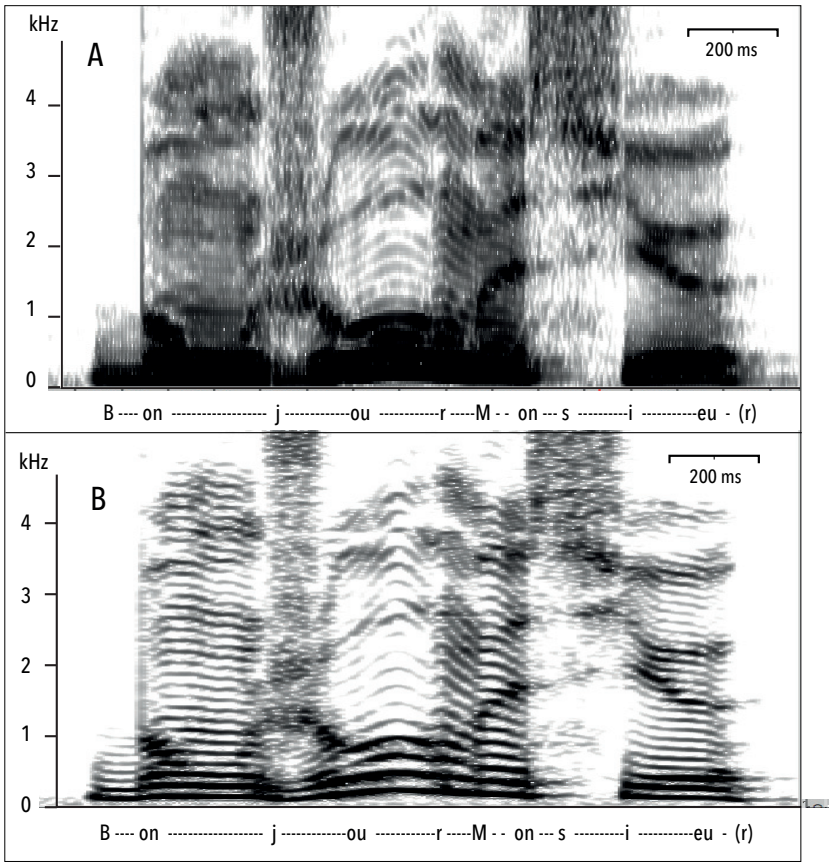


Figure 2.36 Analyse d'une phrase parlée, réalisée avec deux valeurs de la transformée de Fourier : en haut, 512 pts, en bas, 2 048 pts (fréquence d'échantillonnage = 44,1 kHz).

Il est habituel en phonétique de privilégier une bonne résolution temporelle, nécessaire pour l'analyse des bruits consonantiques, généralement très brefs. Ce parti pris favorise aussi la mise en valeur des zones de résonance formantique, bien visibles en A, qui sont caractéristiques des transitions vocaliques et, d'une façon générale, de la forme acoustique globale des mots. Pourtant le choix d'une fenêtre temporelle étroite pénalise l'analyse fine des fréquences et en particulier celle de la fréquence fondamentale qui porte l'intonation. Il faut donc procéder à une autre analyse, avec une fenêtre de plus grande taille, comme en B, où les variations du spectre harmonique traduisent celles de l'intonation. Les deux analyses offrent des informations complémentaires.

Remarque : il faudrait idéalement pouvoir disposer d'une échelle temporelle variable, grossissant les événements transitoires et comprimant les parties stables, afin d'adapter les représentations à la perception humaine, essentiellement mobilisée par « ce qui change dans le temps ».

4.9. L'analyse et l'interprétation perceptive : un exemple trompeur

Son 2.23a (5″)

Oiseau 1

Son 2.23b (18″)

Oiseau 2

Les deux analyses d'un chant d'oiseau (figure 2.37) sont tout à fait identiques à l'œil, alors que les deux écoutes produisent une impression totalement différente (Son **2.23a** et Son **2.23b**). Si l'on examine attentivement les échelles de temps et de fréquence, on lit pour l'analyse A : 4,3 s et 7 kHz ; pour l'analyse B : 17 s et 1700 Hz. La deuxième analyse est due au ralentissement et à la transposition vers le grave du premier chant, mais les paramètres de l'analyse FFT ont été ajustés (512 pts pour Oiseau 1, et 2048 pts pour Oiseau 2) afin d'obtenir la même image. Cet exemple illustre la nécessité d'une lecture attentive des échelles d'analyse pour interpréter correctement les représentations visuelles. Il met aussi l'accent sur l'incidence majeure du paramètre temporel sur la perception sonore.

Observez attentivement les échelles de temps et de fréquence.

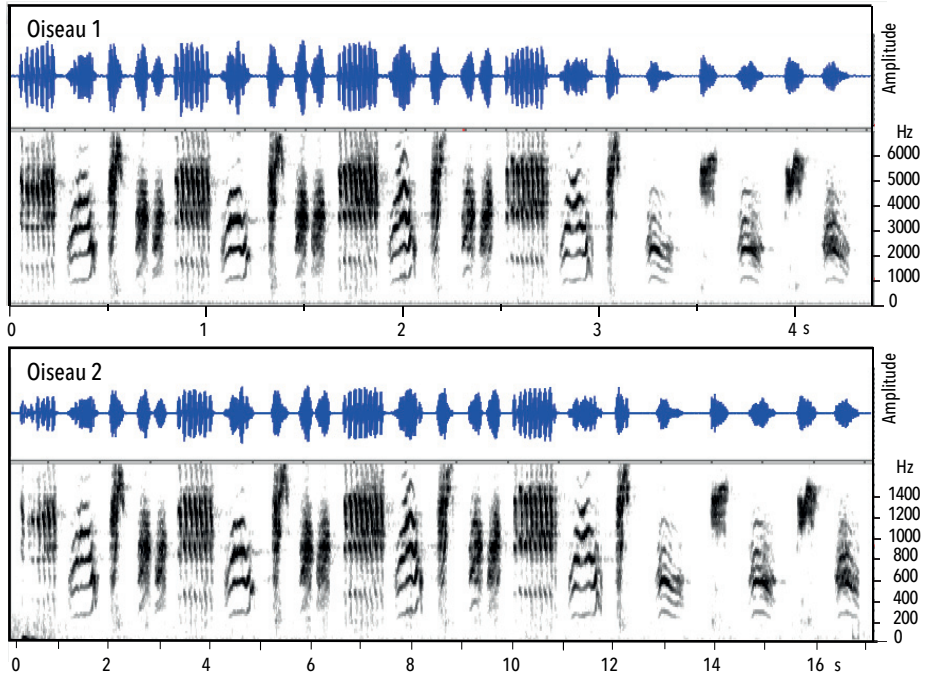


Figure 2.37 Analyse d'un chant d'oiseau : la rousserolle verderolle. Le fragment analysé à la partie supérieure correspond au chant normal entendu dans le Son 2.23a. La partie inférieure de la figure correspond au même fragment ralenti (4 fois) en modifiant la vitesse de lecture de la bande magnétique. Cette opération a aussi pour effet de transposer le son deux octaves plus bas. C'est le Son 2.23b. Malgré l'énorme différence d'écoute entre les deux sons, les analyses apparaissent identiques.

4.10. La voix chantée et le piano : comparaison de sons de hauteur fixe et de sons vibrés

À l'audition (Son **2.24**), pas de doute, c'est bien la même « musique », mais à première vue (figure 2.38) la lecture de ces deux analyses est plutôt déconcertante. Le piano, de type impulsionnel, produit des sons distincts par leurs accents d'intensité (voir la courbe d'amplitude) et par l'impact net du début de chaque frappe visible sous forme de trait vertical sur le sonagramme.

Son 2.24 (16")

Voix chantée puis piano

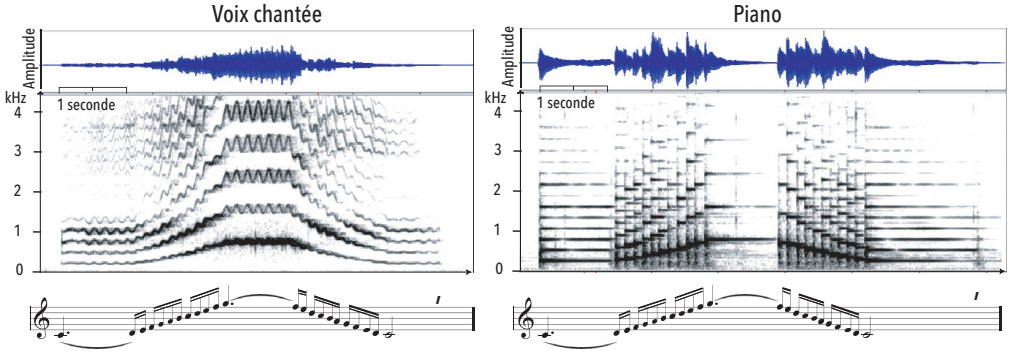


Figure 2.38 Une gamme diatonique chantée et jouée au piano : une même mélodie et deux instruments différents.

Pendant la tenue de la note aiguë (sol_4), le son s'éteint graduellement. À l'opposé, la voix chantée est continue du début à la fin et se développe en intensité sur la note élevée. La différence la plus visible et la plus marquante est celle du tracé des hauteurs. Les notes du piano, de fréquence définie et stable, apparaissent comme autant de colonnes verticales et on peut sans peine y distinguer chacune d'elles. L'analyse de la voix chantée pose au contraire un problème d'interprétation en raison du vibrato, oscillation de part et d'autre de la « note » (voir le chapitre 6, page 268). Cette figure confirme une fois de plus que la représentation du signal acoustique n'est qu'une représentation du phénomène sonore, certes objective, mais qu'il faut apprendre à interpréter pour comprendre l'écoute musicale.

Remarque : l'analyse de la fréquence fondamentale de l'exemple chanté, extraite par autocorrélation, a déjà été présentée figure 2.8, page 42.

4.11. Nuit en Vendée

Au mois de mai, la nuit est habitée : les chants des insectes, batraciens et oiseaux s'entremêlent sans se confondre (Son **2.25**), réalisant ainsi un concert de formes sonores magnifiquement rendues par l'analyse sonagraphique de la figure 2.39. Chaque animal a sa signature et, le plus souvent, chacun occupe une zone du spectre. Comme le concert est dense et fourni, nous avons extrait les sons correspondant aux cadres de la figure en noir et blanc pour les donner à entendre séparément (Son **2.26**). L'écoute de la séquence complète (Son **2.27**) pourra donner lieu à d'autres découvertes.

Son 2.25 (13")

Séquence nocturne (extrait du Son 2.27)

Son 2.26 (19")

Montage des zones : insecte, oiseau 1, oiseau 2 et batracien

Son 2.27 (2'16)

Concert complet avec les vaches et le paon

Son 2.25 (13″) **Son 2.26 (19″)** **Son 2.27 (2′16)**

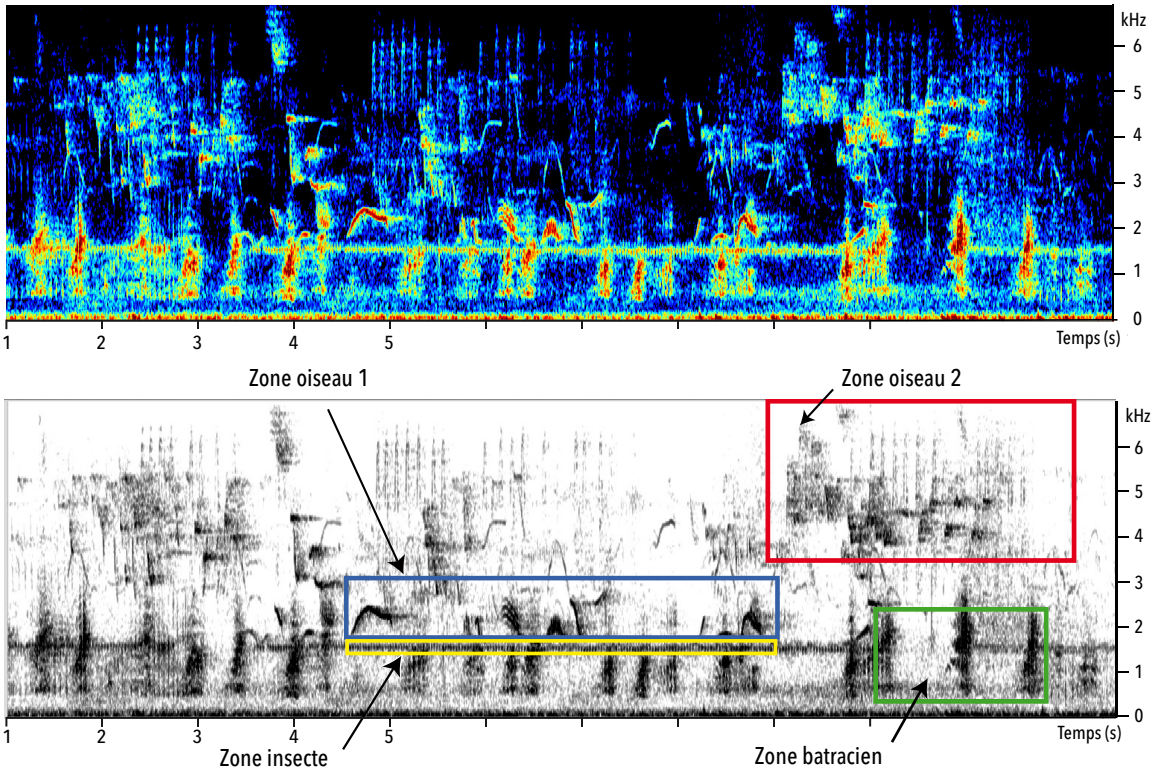


Figure 2.39 Paysage sonore nocturne enregistré au mois de mai. Chaque animal est reconnaissable à sa signature sonore. Les zones spectrales encadrées dans la figure en noir et blanc ont été filtrées et peuvent être écoutées séparément, ce qui aide ensuite à écouter plus finement cette séquence particulièrement dense.

4.12. Une séquence sonore de percussions de hauteur indéterminée

Son 2.28 (10″)

Séquence composite d'instruments à percussion

La représentation spectrotemporelle du sonagramme met en évidence l'originalité acoustique des sons produits par les instruments de percussions dits de hauteur indéterminée (Son **2.28**). Dans cette séquence on repère : 1 : la grosse caisse, 2 : le triangle, 3 : la cabasa, 4 : le tambour de basque, 5 : la caisse claire sans timbre, 6 : les *wind chimes*, 7 : le *wood-block*, 8 : le tambour militaire (roulement).

Chaque instrument produit une forme sonore caractéristique qui se signale par :

- la position dans le spectre : graves (1, 5), médium (7), aigus (2, 3, 4, 6),
- la netteté de l'impact (2, 5, 7),
- l'évolution temporelle : les roulements (4, 8) et la succession des micro-impacts ascendants (6).

Le lecteur peut se reporter au chapitre 7, § 2.2.1, où ces instruments sont étudiés.

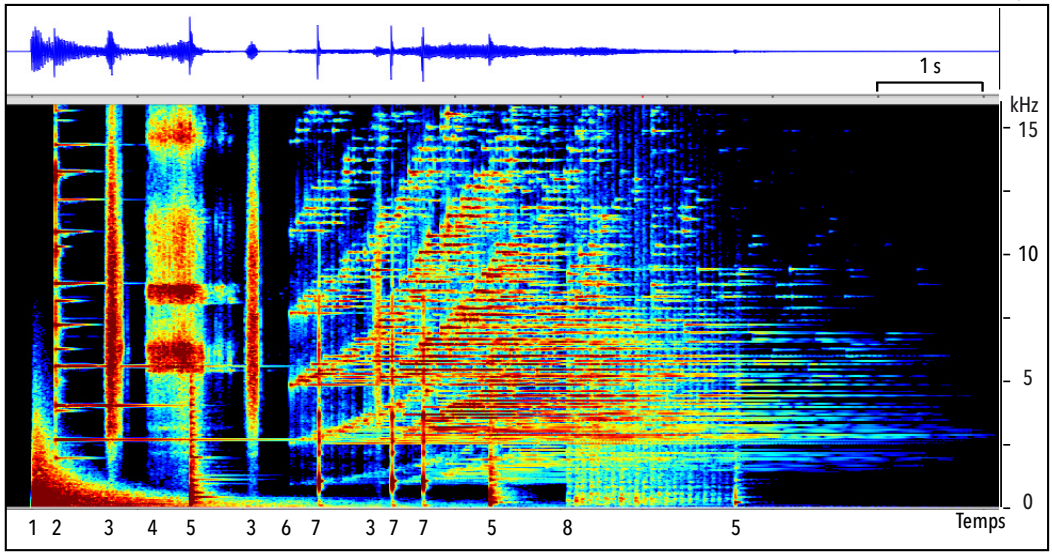


Figure 2.40 Séquence musicale comportant plusieurs instruments de percussion de hauteur indéterminée. La forme sonore de chacun d'eux, « lisible » à l'écoute, est facilement identifiable sur la représentation sonagraphique. Plusieurs instruments couvrent toute l'échelle sonore. Échelle de fréquence : 16 kHz ; FFT : 2 048 pts.

4.13. L'analyse d'un son de hautbois avec trois logiciels différents : Wavesurfer, Praat et Audiosculpt

L'analyse spectrale, accessible à tous, nécessite tout de même un minimum de connaissances acoustiques. Prenons pour exemple trois logiciels disponibles en 2014 : Wavesurfer (WS), Praat (PR) et Audiosculpt (AS)[12].

Lorsqu'on réalise l'analyse spectrale d'un son de hautbois (Son **2.29**, stéréo) avec les réglages par défaut de chaque logiciel, on obtient les trois images de la figure 2.41. Chaque logiciel affiche la forme d'onde (A) et une représentation sonagraphique (B), mais il n'est pas facile de s'y retrouver lorsqu'on est néophyte.

Son 2.29 (2")

Hautbois, note la_3

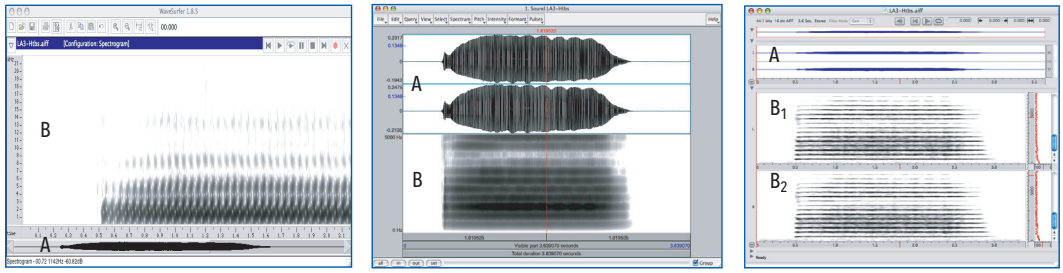
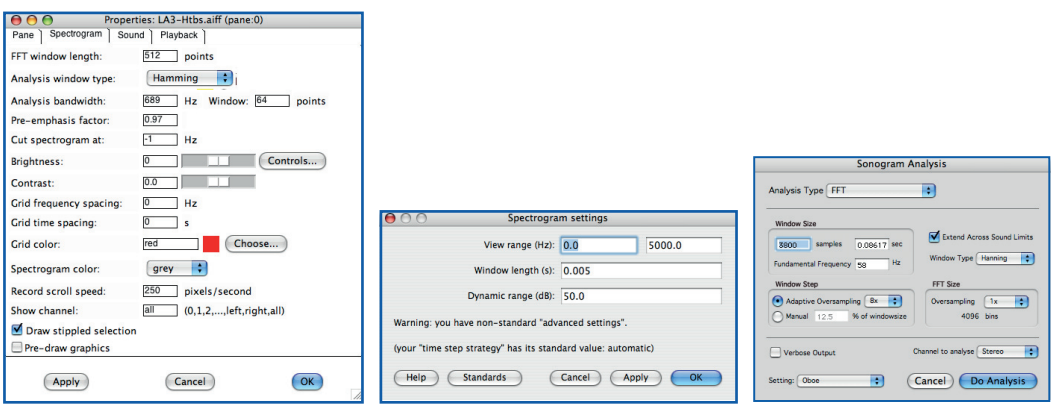


Figure 2.41 Comparaison des images produites par trois logiciels lors de la première ouverture avec les paramètres par défaut.

12. Ce choix correspond à notre expérience : WS offre le grand avantage d'une analyse en temps réel ; PR est remarquablement documenté sur les différentes analyses ; AS permet de transformer les sons de façon intuitive, en travaillant directement la représentation graphique. De nombreux autres logiciels sont disponibles et peuvent être essayés.

Bien que le fichier son soit stéréo, on remarquera que, selon les logiciels, nous avons une seule forme d'onde et un seul sonagramme, deux formes d'onde et un sonagramme ou deux formes d'onde et deux sonagrammes. Quand aux différences d'aspect des sonagrammes, elles tiennent aux choix des échelles et aux réglages des paramètres.

- Échelle de temps (axe horizontal). À l'ouverture, WS n'affiche que 2 s de son. On change la durée d'affichage à l'aide des loupes.
- Échelle de fréquence (axe vertical). Chaque type d'analyse a une valeur limite supérieure en fréquence par défaut qui est particulière : 21,5 kHz pour WS ; 5 kHz pour PR ; 7 kHz pour AS. C'est un point important pour la lecture. Pour modifier cette valeur, il faut accéder aux *settings* (*properties* dans WS) et ajuster la limite supérieure de l'analyse à la valeur désirée pour l'affichage. Avec AS, l'affichage s'ajuste manuellement au moyen de « l'ascenseur » de droite.
- Paramètres de l'analyse de Fourier. Les différences d'aspect de l'analyse spectrographique proviennent principalement des valeurs choisies pour l'analyse, or chaque programme possède des valeurs par défaut (figure 2.42). Ouvrons la fenêtre *properties* de WS et la fenêtre *spectrogram settings* de PR. La taille de la fenêtre est spécifiée soit par le nombre de points (512 pour WS) soit par la durée en seconde (0,005 s pour PR). En ce qui concerne AS nous avons choisi un setting par défaut, *oboe,* dont les réglages apparaissent dans la fenêtre *sonagram analysis* sous deux formes : en nombre d'échantillons (3800) et en secondes (0,08617). La valeur de *fundamental frequency* affichée par ailleurs correspond à celle que peut fournir le programme lorsqu'on demande un calcul automatique de la fréquence fondamentale. Cette valeur vaut 5 fois Δf (qui est 1/Δt).



WS PR AS

Figure 2.42 Paramètres « par défaut » de l'analyse de Fourier tels qu'on les trouve dans les trois logiciels de nos analyses. Successivement Wavesursurfer, Praat et Audiosculpt.

Avec un peu d'habitude, ces valeurs deviennent vite familières. Nous donnons ci-après les correspondances entre le nombre de points, Δt et Δf pour une fréquence d'échantillonnage (*fe*) de 44,1 kHz. Le son de hautbois est un $la_3 = 440$ Hz (période = 0,00227 s).

Tableau 3. Exemples de relations Δf.Δt pour *fe* = 44,1 kHz

Taille de la transformée (points)	Résolution temporelle Δt	Résolution fréquentielle Δf	Limite inférieure d'une bonne détection de f
FFT = 256 pts	Δt = 0,005 s	Δ*f* = 172 Hz	861 Hz
FFT = 512 pts	Δt = 0,011 s	Δ*f* = 86 Hz	430 Hz
FFT = 1024 pts	Δt = 0,023 s	Δ*f* = 43 Hz	215 Hz
FFT = 2048 pts	Δt = 0,046 s	Δ*f* = 21,5 Hz	107,6 Hz
FFT = 4096 pts	Δt = 0,092 s	Δ*f* = 10,7 Hz	53 Hz

5. En guise de conclusion

L'analyse acoustique met à notre disposition des outils pour explorer le contenu physique des sons. En tenant compte de toutes les précautions prises au moment de l'acquisition du son (qualité et positionnement des capteurs), une analyse fournit déjà un support objectif : c'est le point de départ d'une étude de la perception des sons produits par les instruments de musique.

Il n'existe pas d'analyse idéalement bonne, mais des analyses diversement adaptées aux objectifs que l'on se donne. En acoustique musicale, l'oreille est le meilleur guide. C'est elle qui oriente le choix du type d'analyse à mettre en œuvre, la sélection des paramètres pertinents et souvent la découverte de phénomènes que l'écoute habituelle occulte.

Les outils numériques disponibles aujourd'hui ont totalement transformé notre rapport à l'analyse des sons. En particulier l'analyse en temps réel, c'est-à-dire la possibilité de voir une représentation à l'instant même[13] où le son se produit, permet de repérer dans l'image les parties qui changent significativement dans le temps de l'écoute. Dans un second temps, l'observation des analyses mémorisées donne accès à des transformations qui permettent de tester la validité de l'interprétation. Pour un auditeur doué d'une bonne oreille, l'analyse réalise une amplification de l'écoute grâce au couplage entre la vision et des écoutes modifiées – en particulier l'allongement de la durée d'un son (indépendamment de la fréquence), la réjection de certaines parties, la sélection et l'écoute de composantes particulières. Il ne faut pas hésiter à multiplier les représentations d'un son donné, changer les échelles de fréquence et de temps, les valeurs des paramètres de l'analyse FFT, les caractéristiques de l'image visuelle (contraste, couleurs), jusqu'à obtenir la représentation la plus lisible. De tels documents peuvent être précieux :

- pour le chercheur en cognition qui désire travailler sur les sons « naturels, réels », car bien des idées acquises en perception sonore sont issues d'expériences pratiquées avec des sons conçus spécialement en laboratoire, alors que nous développons notre expérience de l'écoute sur la base des sons complexes de notre environnement : parole, bruits divers et sons de la musique ;
- pour le luthier confronté aux avis contradictoires des auditeurs, qui souhaite prendre du recul par rapport à ses habitudes d'écoute personnelle, et qui souvent se construit un système d'interprétation de la perception pouvant être en porte-à-faux par rapport à la réalité physique du monde sonore ;
- pour le compositeur désireux de prendre en compte la structure spectrale des sons instrumentaux dans son écriture. Plusieurs compositeurs comme Gérard Grisey, Jonathan Harvey et Tristan Murail se sont directement servis de l'analyse acoustique dans leur œuvre, comme point de départ ou comme grille d'écriture.

13. Il subsiste bien sûr un délai entre la production du son et son analyse, mais il est devenu imperceptible.

Bref historique de l'analyse acoustique

Les progrès réalisés dans la connaissance des phénomènes vibratoires, et plus particulièrement des ondes sonores, sont étroitement liés aux avancées des techniques d'observation des vibrations sonores. L'ouïe, la vue et le toucher sont les premiers « outils » d'investigation. Dans l'histoire de l'acoustique, deux notions se sont rapidement établies : la hauteur (grave/aigu) du son, liée à la vitesse des mouvements vibratoires ou « tremblements », et la force du son (intensité) liée à son amplitude. Le timbre que connaissent bien les musiciens est longtemps resté un phénomène mystérieux. En 1636, le père Marin Mersenne note « la présence de la « douzième et de la dix-septième » dans un son de viole de gambe. « Il faut [dit-il] écouter dans une pièce très calme, de préférence la nuit[1]. » Cependant, les auteurs de cette époque ne font pas du tout la relation entre la perception de tels phénomènes et le timbre. C'est le géomètre Monge qui aurait le premier donné une explication du timbre, la « coexistence des petites oscillations », phénomène physique dont l'outil d'analyse mathématique ne sera fourni par J. B. Fourier qu'en 1822 (*Théorie analytique de la chaleur*, Didot, Paris).

L'idée qu'un mouvement vibratoire complexe puisse être considéré comme la résultante de différents mouvements primaires va se développer au cours du XIXe siècle, avec les progrès de l'observation expérimentale de ces mouvements.

Les premières analyses systématiques sont faites à l'oreille par Helmholtz à l'aide d'une série de résonateurs[2] spécialement construits à cet effet. Les efforts vont alors se porter sur la matérialisation du son, afin de rendre visible la forme temporelle de la vibration sonore. Ce sont d'abord les flammes « sensibles » (Tyndall, 1880) et les tracés sur tambours enduits de noir de fumée.

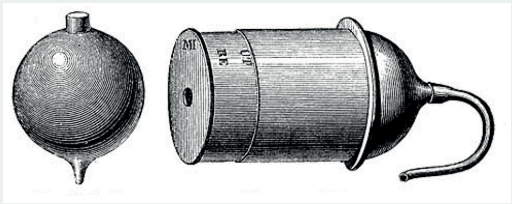


Figure 2.43 Résonateurs de Helmholtz. Le résonateur cylindrique est réglable grâce à l'extrémité coulissante.

Koenig, R., 1889, figures 15 et 16 (Document Virtual Laboratory).

En 1889, Rudolph Koenig construit un analyseur combinant les résonateurs de Helmholtz et les flammes sensibles. Chaque résonateur est relié à un bec de gaz dont la flamme varie en proportion de l'amplitude vibratoire de l'air dans le résonateur. Grâce au miroir tournant à quatre faces qui effectue une sorte de stroboscopie manuelle, on peut obtenir une image continue des variations temporelles de la hauteur des flammes. Ces images pouvaient être photographiées pour être ensuite étudiées. Il devient possible de visualiser la composition des vibrations et leurs amplitudes relatives (voir figure 2.44).

Ces appareils, d'une grande ingéniosité, ont joué un rôle important dans la démonstration des phénomènes acoustiques. Ils nécessitent de produire le son à chaque nouvelle analyse, puisque le travail se fait en direct.

L'enregistrement durable qui « mémorise » la vibration, d'abord sous forme d'un tracé (phonautographe de Scott de Martinville, 1857) puis par déformation d'une matière malléable (cire), ce qui va permettre de le reproduire ensuite (Charles Cros et Edison, 1877), donne enfin accès à une analyse objective. Une nouvelle technique d'analyse se développe avec l'observation au microscope du « phonogramme » ou tracé de l'inscription du son en profondeur, dans la cire du cylindre enregistreur[3]. Le son n'existe plus seulement dans la mémoire des auditeurs. Il peut être reproduit à volonté, et peut faire l'objet d'études objectives.

À la fin du XIXe siècle, la transformation de la vibration sonore en courant électrique est une étape décisive qui donne accès à la mesure (en volts). L'amplification du signal sonore permet, en outre, une meilleure restitution de la gamme des fréquences et l'étude de phénomènes de très faible amplitude auxquels l'oreille est très sensible. Avec l'apparition des microphones, de l'oscillographe

1. Mersenne, 1636, Livre quatrième des instruments, Proposition IX.
2. Un résonateur est ajusté pour une fréquence précise. Il faut donc en construire autant que le nombre des harmoniques que l'on désire analyser, et recommencer l'opération pour un son d'une autre fréquence ! D'où l'idée de réaliser des résonateurs réglables en volume, comme celui de la figure 2.43.
3. Voir l'ouvrage de Gellé (1899), *L'audition et ses organes*, qui utilise exhaustivement les tracés du phonogramme pour appuyer son analyse des sons.

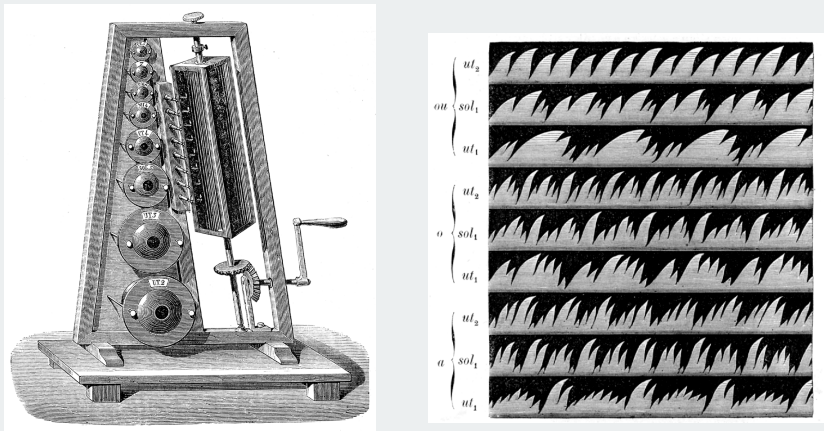


Figure 2.44 Analyseur de Kœnig (1889). À gauche : appareil construit pour l'analyse des huit premiers harmoniques d'un son stable. Celle-ci est matérialisée par les modifications des flammes manométriques sensibles aux variations de pression dans les résonateurs. À droite : l'image du mouvement des flammes qui caractérisent les voyelles « a, o, ou » chantées chacune sur les trois notes ut_1, sol_1, ut_2.

Tyndall, J., 1869, figures 168 et 169.

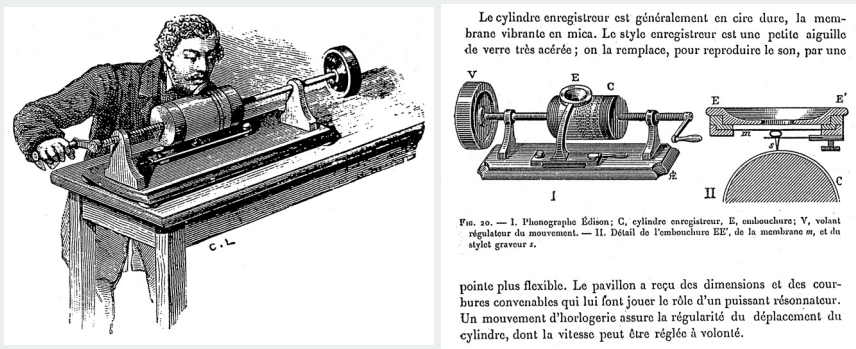


Figure 2.45 Principe de l'enregistrement sur cylindre. Pour obtenir une gravure suffisante, il faut s'approcher très près du pavillon au fond duquel se trouve la membrane réceptrice, munie d'une pointe s'enfonçant dans la cire.

À gauche : Radau, R., 1880, figure 81. À droite : source non identifiée.

cathodique, des circuits résonants (filtres) puis du ruban magnétique, on peut enfin étudier la forme d'onde, mesurer l'amplitude et la fréquence d'un signal périodique, visualiser les transitoires, procéder à l'analyse spectrale.

La digitalisation du signal sonore (voir Glossaire) représente une autre étape importante. Elle donne accès à une grande précision des mesures, et surtout permet une synthèse sonore finement contrôlée : c'est l'outil privilégié pour l'exploration de la perception auditive.

Pendant plus d'un siècle, seuls les sons périodiques stables ont pu être analysés. Le développement des techniques mathématiques de traitement du signal, comme la FFT, l'analyse par ondelettes ou l'analyse Wigner-Ville, permettent aujourd'hui de contourner les difficultés posées par les signaux instationnaires. Toutefois, un problème de fond subsiste, lié à la définition même de la fréquence. Plus les signaux sont brefs, plus la mesure des fréquences est incertaine. Pour contourner cette difficulté, plusieurs stratégies d'analyse existent. Voir § 3.3, page 43.

Curieusement, le courant de la musique concrète initié par Pierre Schaeffer dans les années 1950 s'est détourné pendant longtemps de l'analyse acoustique des sons, alors même que la complexité des signaux et de leurs transformations rendait obsolète toute notation traditionnelle. Il est vrai qu'à cette époque le sonagraphe analogique, seul outil offrant une représentation spectrotemporelle significative, ne permettait d'analyser que 2,4 s de son – ce qui est bien peu pour un musicien – et qu'il fallait attendre 10 minutes avant d'obtenir le résultat. Le projet d'un « solfège concret » devant fournir une description exhaustive des matériaux sonores de la nouvelle musiques est resté à l'état de grille théorique. Soixante ans plus tard, riches de l'élargissement de l'écoute par la diversité des musiques expérimentales et des musiques traditionnelles, munis de toutes les possibilités de contrôle des sons pour expérimenter la perception, nous sommes devenus beaucoup plus modestes quant à l'interprétation perceptive des analyses physiques du son. Le bon sens nous conduit à écrire qu'une analyse est pertinente en acoustique musicale lorsqu'elle fournit une représentation interprétable des caractéristiques perçues.

6. Les sons du chapitre 2

Son 2.1 – Exemple de quatre sons produits avec la bouche : un sifflement, un chuintement, une voyelle chantée en glissando ascendant et deux claquements de langue. [M. C.]

Son 2.2 – Gamme diatonique ascendante chantée de do_3 à sol_4, par une soprano : C. Herzog ; 1979. [Archives LAM]

Son 2.3 – Cor des alpes. Signal recueilli au pavillon d'un cor des Alpes qu'on excite par une impulsion de la main sur l'embouchure. Fréquence 46,72 Hz (environ fa_0). Voir son 1.23. [M. C.]

Son 2.4 – Son grave joué au tuba : do_1 ± 65,4 Hz ; fichier son « Tuba-Iowa-C0-ff » [Sound Database Iowa]

Son 2.5 – La note do_4 jouée avec deux instruments : un vibraphone et un tuyau d'orgue. [M. C.]

Son 2.6 – Variations de l'amplitude sonore dues au rayonnement. Au début, son d'un verre frotté par rotation du doigt sur le bord, ce qui entraîne une variation angulaire du rayonnement. Ensuite son d'un petit gong birman tournant verticalement après la frappe. [M. C.]

Son 2.7 – Une gamme chromatique sur toute la tessiture est jouée sur la flûte traversière puis sur le hautbois. Voir l'analyse moyennée du contenu spectral. [Archives LAM]

Son 2.8 – Comparaison de la qualité sonore de deux jeux d'un orgue, après (1999) et avant (1997) restauration ; mêmes registrations et mêmes conditions d'enregistrement. Son 2.8a, annonce : « *Nivers, Grand Fond d'orgue, avec la quarte de nasard, version courte* ». Son 2.8b, annonce : « *fourniture 3 rangs* » puis jeu d'une gamme diatonique, de do_3 à sol_4, sur la fourniture du Positif. Le diapason de l'instrument est bas (la_3 = 392 Hz). Orgue d'Ebersmunster, mission du ministère de la Culture. [Archives LAM]

Son 2.9 – Étude du transitoire d'attaque d'un tuyau d'orgue grâce à la transformation d'un son réel par synthèse sonore. On entend quatre exemples, chacun étant répété deux fois. Les lettres renvoient à la figure 2.21. 1/ Tuyau d'orgue original (A). 2/ Tuyau privé du son de bouche du transitoire par filtrage (C). 3/ Son

de bouche isolé (D). 4/ Le son de bouche isolé (D) suivi du son du tuyau original (A). [M. C.]

Son 2.10 – Une mélodie jouée par un son pur (sinusoïdal) qui est en réalité le premier harmonique filtré d'une flûte traversière (voir Son 2.11). [M. C.]

Son 2.11 – Mélodie jouée à la flûte traversière ; Benoît Fabre. [LAM]

Son 2.12 – Court exemple vocal d'un grand intervalle chanté avec vibrato. Vous pouvez écouter la phrase complète au chapitre 6, Son 6.57. [Archives LAM]

Son 2.13 – Synthèse d'un glissando ascendant/descendant, avec trois formes d'ondes. Son 2.13a : onde sinusoïdale, onde rectangulaire, onde carrée. Voir la figure 2.27 pour l'analyse. Le Son 2.13b permet de comparer rapidement les trois types de sons sur une même note. [M. C.]

Son 2.14 – Un son de steel-drum, instrument à percussion des îles Caraïbes. [P. Gaillard, 2000]

Son 2.15 – Son d'un verre de cristal[14] percuté et entretenu sur les modes 1 et 2. Successivement : son percuté, entretien du partiel 1 (fa_4) ; entretien du partiel 2 (mib_5). Immédiatement après, réécoutez l'exemple et vous distinguerez nettement le son du mib_5 dans son du verre percuté. [M. C.]

Son 2.16 – Octaviation et quintoiement : séquence courte. Les deux premiers partiels d'une flûte et d'une clarinette. [Voir Son 1.17]

Son 2.17 – Les partiels de la flûte traversière en ut, doigté do_3, tous les trous bouchés. Succession de huit sons dits « harmoniques », qui sont en réalité les partiels des modes propres du tuyau. [Joe Wolfe, site Internet]

Son 2.18 – Les partiels de la clarinette en *sib*, doigté tout bouché. Succession de six sons dits « harmoniques », ou partiels des modes propres du tuyau. [Joe Wolfe, site Internet]

Son 2.19 – Tilinca. [Voir Son 1.25]

Son 2.20 – Partiels du tuba. Le fondamental est fa_0. Jeu des dix premiers partiels (ascendant puis descendant) ; G. Bucquet [Voir Son 1.15]

Son 2.21 – Exemple didactique de chant harmonique (ou diphonique) : glissando spectral ascendant et descendant original puis transposé. Trân Quang Hai. [Voir Son 1.12]

Son 2.22 – La phrase « Bonjour monsieur », dite par une voix d'homme. S. Lamesch. [Archives LAM]

Son 2.23 – La rousserolle verderolle (*Acrocephalus palustris)*. Son 2.23a : bref extrait du chant « normal ». Son 2.23b : le même extrait ralenti 4 fois et transposé deux octaves plus bas devient méconnaissable et prend des accents humains (transposition sur bande analogique). [Archives LAM]

Son 2.24 – Comparaison de la gamme diatonique do_3-sol_4 chantée puis jouée au piano. Chant : Colette Herzog, 1970. [Archives LAM]

Son 2.25 – Concert nocturne. Extrait correspondant à la figure 2.39. [M. C.]

Son 2.26 – Concert nocturne. Montage des « zones » spectrales filtrées de la figure 2.39. Successivement : stridulation d'insecte, oiseau 1, oiseau 2, batracien. [M. C.]

Son 2.27 – Concert nocturne complet, pour le plaisir d'écoute. Enregistré au lieu-dit de l'abbaye d'Orouet, Vendée, 1982. [M. C.]

Son 2.28 – Séquence musicale réalisée avec des sons d'instruments à percussion de hauteur indéterminée. Sons 1 à 8 extraits de la base de données RWC : grosse caisse, 411BD1N3 ; triangle, 411TANO3 ; cabasa, 445CANO3 ; tambourin

14. Verre différent de celui qui est analysé au chapitre 1 (sons 1.1 ; 1.2 ; 1.7 ; 1.8).

basque, 445TMROF ; caisse claire sans timbre, 411SD1N3 ; cabasa, 445CANO3 ; *wind chimes*, 445WCBU3 ; *wood block*, 411WBNO3. Son 9 extrait de la base McGill : tambour militaire, 2e roulement à la brosse, Vol. 3, index 11-06. [M. C.]

Son 2.29 – Hautbois : note la_3. [M. C.]

CHAPITRE 3

LE SYSTÈME AUDITIF HUMAIN

1. Description sommaire de l'organe auditif

1.1. Introduction

Comme tous nos organes des sens, l'oreille renferme des éléments capables de capter les signaux du monde extérieur et de les transmettre à une structure spécialisée, ici la cochlée, qui réalise la conversion des mouvements mécaniques en influx nerveux. Elle nous renseigne sur les événements se produisant hors du champ visuel ou à grande distance, pour autant que l'énergie transmise soit suffisante pour déclencher la sensation auditive, et que les vibrations soient dans la gamme de fréquences audibles. L'oreille capte aussi directement les vibrations se produisant à l'intérieur du corps comme le grincement des dents, la déglutition, les battements du cœur. Enfin, l'oreille interne comprend également deux organes nous informant, l'un sur la position statique de notre corps, l'autre sur les déplacements qui affectent le corps dans les différentes directions de l'espace en captant les variations de la vitesse.

Le système auditif dans son ensemble est donc un capteur d'énergie mécanique qui nous renseigne de façon fine sur les différents mouvements externes ou internes à notre corps. Ce n'est qu'au niveau des centres nerveux supérieurs que les signaux sonores et vibratoires prennent sens. La parole, la musique, les bruits de l'environnement sont des catégories que chaque sujet élabore et réorganise sans cesse au cours de son expérience personnelle.

1.2. Schéma général du système auditif

Depuis l'entrée du conduit auditif jusqu'au cerveau, le système auditif constitue une chaîne complexe comprenant un capteur, l'oreille proprement dite, un réseau de transmission codée réalisant un prétraitement des signaux et un centre de traitement de l'information sonore nécessitant divers types de mémoires ainsi que des structures d'analyse temporelle des signaux.

La simple perception d'un son implique le bon déroulement d'une succession précise d'opérations.

Une onde aérienne parvient au pavillon de l'oreille, entre dans le conduit auditif, fait vibrer le tympan auquel est fixée la chaîne des osselets qui vibre à son tour et transmet ses vibrations aux liquides de l'oreille interne. À l'intérieur de l'oreille interne, le déplacement des liquides entraîne la déformation de membranes qui supportent les cellules sensorielles. Celles-ci, munies de cils rigides, déclenchent

la production d'influx nerveux qui progressent dans le nerf acoustique puis atteignent une succession de centres nerveux relais (les noyaux), avant de parvenir au cortex cérébral.

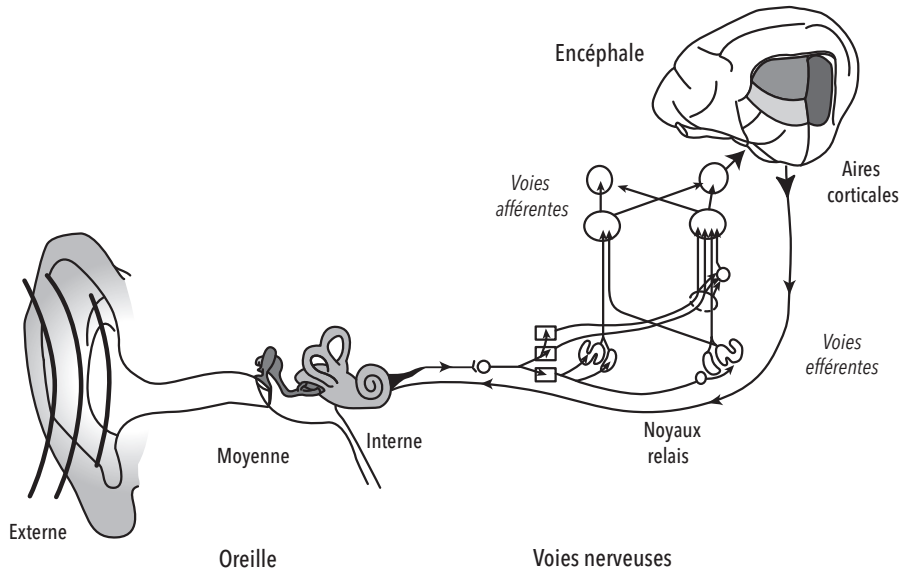


Figure 3.1 Schéma du système auditif humain. Depuis le pavillon jusqu'à l'oreille interne, les vibrations sonores subissent une série de transformations analogiques. À la sortie de l'oreille interne, des impulsions codées transitent dans une suite de centres nerveux avant de parvenir aux aires auditives du cerveau. Des voies nerveuses descendantes, ou efférentes, véhiculent aussi des informations issues des centres supérieurs vers les organes de l'oreille dont elles peuvent modifier le fonctionnement.

D'emblée, nous saisissons deux parties distinctes dans cette chaîne (figure 3.1). La première partie, qui concerne la transmission des vibrations jusqu'aux cellules de l'oreille interne, aboutit à la conversion des ondes vibratoires mécaniques en impulsions nerveuses : c'est l'oreille proprement dite, ou système auditif périphérique. La deuxième partie concerne le transit des influx nerveux dans différents centres où se produisent divers traitements des signaux et en particulier la comparaison avec ceux de l'oreille opposée ou avec ceux d'autres organes sensoriels.

L'ensemble de l'oreille, des voies et des centres nerveux constitue le système auditif.

Le fonctionnement mécanique de l'oreille externe et moyenne est connu depuis longtemps. Il a inspiré les inventions du phonautographe de Scott de Martinville et du téléphone de Graham Bell.

La question qui intrigue le plus les chercheurs est celle de l'analyse des sons et, plus particulièrement, de la séparation des fréquences. On doit à Helmholtz, qui le premier a effectué une analyse harmonique des sons complexes au moyen de résonateurs, l'idée qu'une analyse similaire pouvait se produire dans l'oreille interne. Il fait jouer ce rôle aux fibres de la membrane basilaire en les comparant à des cordes qui vibreraient « sous influence », c'est-à-dire par sympathie. Bien plus tard (1960),

Békésy a montré expérimentalement que cette membrane vibrait effectivement de façon différenciée selon la fréquence.

L'étude de la deuxième partie du système auditif, celle du codage des données sensorielles et des modes de traitement des signaux neuronaux, n'a pu se développer qu'avec les progrès de l'électrophysiologie et de la neurobiologie. On sait maintenant qu'il existe, outre les voies nerveuses ascendantes (afférentes), qui transmettent les données provenant de l'oreille vers les centres nerveux, des voies descendantes (efférentes) en provenance des centres supérieurs, qui contrôlent en partie le fonctionnement de l'organe sensoriel.

Une grande partie de nos connaissances sur le fonctionnement du système auditif provient des résultats d'expériences réalisées avec des animaux de laboratoire soumis à l'audition de sons sinusoïdaux : elle constitue la base physiologique commune à tous les mammifères. Les problèmes deviennent plus complexes lorsqu'il s'agit d'interpréter les résultats de tests effectués avec des sujets humains, dont on sait que les réponses peuvent varier selon le protocole. Pour parer à ces aléas, des méthodes rigoureuses ont été élaborées afin d'explorer systématiquement les réponses des sujets aux variations des trois paramètres d'un son sinusoïdal (la fréquence, l'intensité, la durée) dans des conditions strictes de présentation : sons isolés, sons simultanés, en présence ou non de bruits. Ces recherches sont actuellement stimulées par le développement des implants cochléaires, prothèses excitant directement le nerf auditif de patients dont les cellules nerveuses de l'oreille interne ne sont plus fonctionnelles.

Le corpus des connaissances de la psychoacoustique et de la neurobiologie, bien que considérable, peut laisser le lecteur insatisfait. Même si les stimuli employés sont de plus en plus complexes, ils restent dépourvus de sens pour l'auditeur et, en tant que tels, fort éloignés des sons réels qui nous concernent (voir le chapitre 4). C'est pourquoi les chercheurs placent aujourd'hui beaucoup d'espoir dans les techniques d'imagerie cérébrale qui permettent d'observer l'activité des centres nerveux supérieurs de sujets auxquels on fait écouter de la parole ou de la musique. Il faut bien dire que les résultats sont encore fragmentaires, car les expériences sont lourdes, complexes et coûteuses, et leur interprétation assez dépendante de la variabilité des sujets. En effet, l'activité nerveuse observable « en bout de chaîne » rend compte non seulement du traitement neurophysiologique des données, mais aussi des comportements individuels d'écoute développés tout au long de la vie.

L'étude de l'audition est en plein essor, donc en constant renouvellement. Les notions présentées dans ce chapitre sont volontairement sommaires mais doivent permettre à ceux qui le souhaitent d'aborder les ouvrages spécialisés cités en référence.

Examinons tout d'abord la structure et les particularités du capteur oreille qui transforme des variations de pression sonore en impulsions nerveuses.

1.3. Les trois parties de l'oreille

Du point de vue fonctionnel, l'oreille est constituée de trois parties distinctes (figure 3.2) :

- *l'oreille externe,* qui recueille et conduit les vibrations aériennes jusqu'à la membrane du tympan ;
- *l'oreille moyenne,* qui transmet les vibrations sonores depuis le tympan jusqu'à la fenêtre ovale, par le moyen de trois petits osselets articulés ;
- *l'oreille interne,* qui convertit les vibrations acoustiques en impulsions nerveuses.

À chaque étape, l'onde vibratoire change de milieu de propagation, passant d'un milieu aérien, à un milieu solide (os), puis liquide (périlymphe).

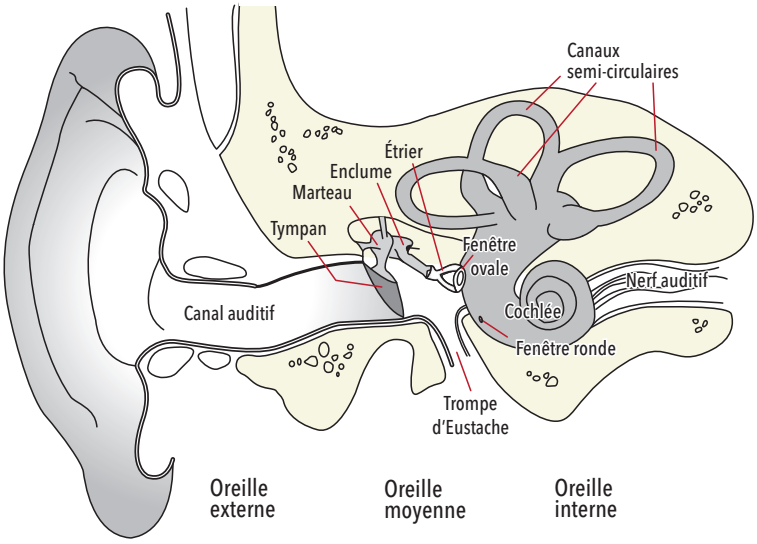


Figure 3.2 Vue générale de la structure du système auditif périphérique.
D'après Lindsay, P., Norman, D., 1980, fig 4.2.

1.3.1. L'oreille externe : un capteur directionnel

Elle comprend le pavillon, le conduit auditif externe et la membrane du tympan.

Le pavillon de l'oreille humaine, fixé à la tête[1], joue un rôle complexe dans la *directionnalité de l'écoute*. Lorsque surgit un son imprévu, nous orientons de façon réflexe la tête dans sa direction (voir page 127). Mais s'il est difficile à entendre, soit parce qu'il est faible, soit parce qu'il est très aigu, nous orientons plutôt l'une des deux oreilles dans la direction du son, en cherchant l'incidence la plus favorable à sa réception.

Pour mesurer la directionnalité de l'oreille, on fait l'expérience suivante : un microphone miniature étant placé au niveau du tympan, une suite de sons sinusoïdaux de fréquence variable et d'intensité constante est présentée selon différents angles, dans un plan horizontal passant par le conduit auditif. Les courbes de variation d'intensité en fonction de la fréquence ainsi tracées rendent compte de la géométrie combinée du pavillon et du conduit.

1. À la différence de nombreuses espèces animales qui peuvent orienter leur pavillon de façon indépendante pour chaque oreille.

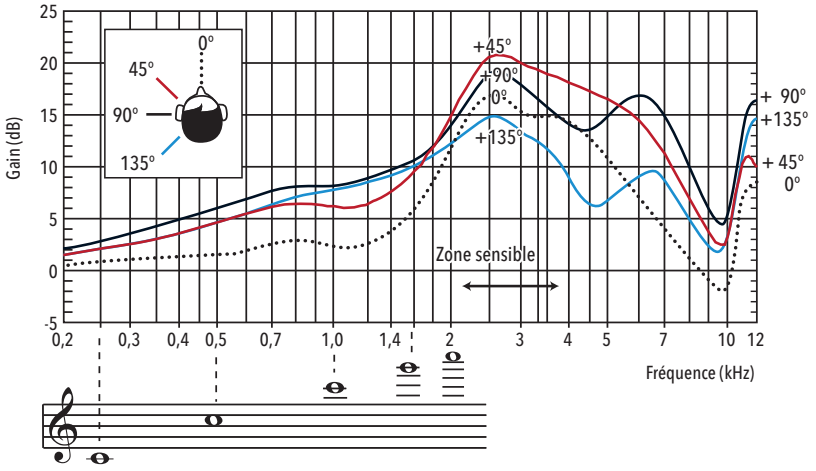


Figure 3.3 Directivité. "Gain" moyen en pression de l'oreille externe en fonction de l'angle d'incidence du son. Mesures dans le plan horizontal passant par les tympans. La sensibilité est maximale pour un angle de 45° ; elle change notablement avec la fréquence.

D'après Pickles, J., 1982, figure 2.2 (selon Shaw, 1974, figure 5).

La figure 3.3 montre les résultats pour quatre angles d'incidence par rapport à la face. Pour toutes les directions, on observe une accentuation marquée de la zone comprise entre 2 et 4 kHz, qui provient de la résonance du conduit auditif : c'est la zone de plus grande sensibilité de l'oreille humaine. On remarque ensuite que les courbes ne se suivent pas parallèlement, ce qui signifie que la meilleure incidence de l'oreille par rapport au son change selon la fréquence. Dans l'aigu, entre 5 et 7 kHz, les écarts entre les courbes à 45° et 135° atteignent 10 dB.

Bien évidemment, ces courbes changent selon les individus. Elles font aujourd'hui l'objet de relevés systématiques[2], car de nombreux chercheurs pensent que la « carte spectrale » ainsi associée à chaque personne joue un rôle important dans l'appréciation de l'espace. Par ailleurs, les petits mouvements de tête, que nous effectuons plus ou moins consciemment, nous fournissent des informations sur la complexité du champ sonore environnant et peuvent expliquer, entre autres, le plaisir sur l'on prend à l'écoute des instruments réels dont le rayonnement varie sans cesse et de façon plus complexe que lors de la restitution par enceintes.

Notons enfin une fonction du pavillon de l'oreille mentionnée par E. Leipp (1977, p. 54), mais souvent négligée : celle de *coupe-vent*, qu'il faut attribuer aux contours sinueux du pavillon. Chacun de nous peut en faire l'expérience : les bruits d'écoulement du vent provenant de face s'atténuent considérablement lorsqu'on présente l'oreille « au vent ».

Les fonctions de l'oreille externe sont passives et ne relèvent que de l'anatomie.

1.3.2. L'oreille moyenne et l'adaptation de la transmission sonore

Trois petits osselets, maintenus en suspension dans la cavité de l'oreille moyenne par des ligaments et des muscles, assurent la connexion entre la membrane du tympan et celle de la fenêtre ovale. Ce sont respectivement le marteau, dont le manche est inséré dans la membrane du tympan, l'enclume, et l'étrier solidaire de la fenêtre ovale (figure 3.4). L'ensemble assure trois fonctions importantes.

2. Pour l'établissement des courbes dites HRTF *(Head Related Transfer Function)*.

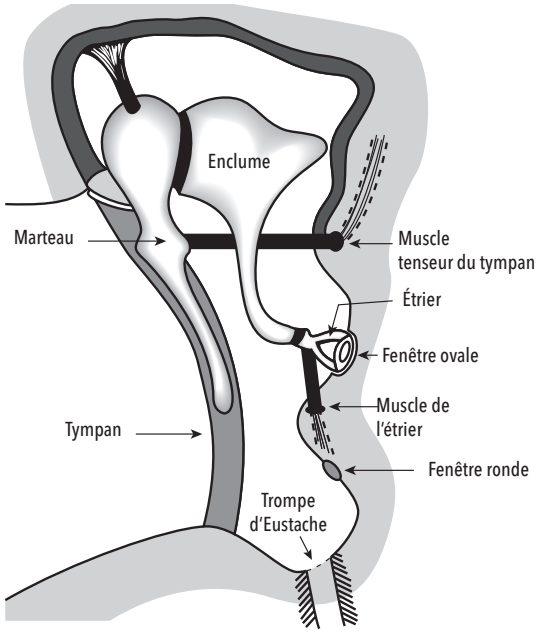


Figure 3.4 Schéma de l'oreille moyenne. Le système des osselets assure une protection du système auditif analogue à celle de l'iris de l'œil. Il est représenté ici, vu d'en haut. Pour la clarté, on a supprimé les tendons, organes passifs, ne servant qu'à maintenir les osselets. Mais on a laissé les deux muscles, organes actifs, ce qui permet de comprendre les mécanismes en présence.

D'après Leipp, E., 1977, *La machine à écouter*, fig. 20.

- La première fonction est une remarquable *adaptation mécanique* des déplacements vibratoires assurant une transmission efficace malgré le changement de milieu de transmission. Il faut, pour déplacer une tranche de liquide (côté fenêtre ovale), une force d'application plus importante que pour déplacer une tranche d'air (côté tympan). L'oreille moyenne réalise cette adaptation par le jeu des rapports de surface vibrante et celui des bras de levier de la chaîne des osselets.
- La deuxième fonction, liée à la première, consiste dans le *réglage de la plage dynamique*, c'est-à-dire l'intervalle entre les sons les plus faibles et les sons les plus forts. Deux muscles – le muscle tenseur du tympan et le muscle de l'étrier – peuvent se contracter en bloquant partiellement les mouvements de la chaîne des osselets, réalisant ainsi une compression de l'ordre de 20 dB (Pickles, p. 22), donc une protection de l'oreille interne contre les sons trop intenses. La contraction des muscles se met en jeu de façon réflexe lorsque nous parlons ou chantons, et peut être volontairement activée en prévision d'un bruit intense. Ainsi bloquée, la chaîne des osselets est aussi plus raide : elle transmet moins bien les basses fréquences, ce qui affecte l'équilibre spectral. L'adaptation réalisée par l'oreille moyenne agit donc autant sur l'intensité du signal que sur son spectre.
- L'oreille moyenne assure une troisième fonction, celle de l'*équilibration de la pression aérienne sur les deux faces du tympan*. En effet, pour qu'il puisse répondre à la plus légère surpression sonore provenant de l'extérieur, il faut

que le tympan soit en équilibre, c'est-à-dire que la pression aérienne dans la caisse de l'oreille moyenne soit la même qu'à l'extérieur. La *trompe d'Eustache* qui débouche dans la cavité buccale, assure cet équilibre. D'ordinaire, nous n'en avons pas conscience, puisque elle s'ouvre à chaque déglutition. Mais un rhume (trompe d'Eustache bouchée) ou une dépressurisation rapide (en avion) peuvent provoquer des distorsions auditives que l'on fait disparaître en bâillant activement.

1.3.3. L'oreille interne : le système vestibulaire et la cochlée

La chaîne des osselets aboutit à la fenêtre ovale qui est l'entrée d'une cavité de forme complexe, au nom évocateur de *labyrinthe*, logée dans l'os du rocher (figure 3.5).

On y trouve deux organes intimement connectés : le système vestibulaire et la cochlée (figure 3.6, page suivante).

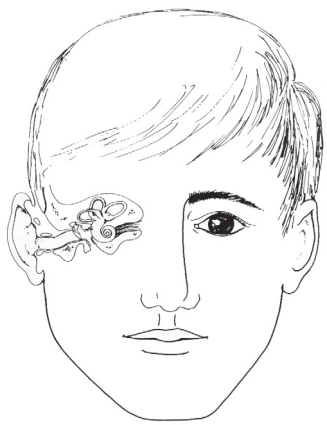


Figure 3.5 Position des organes du labyrinthe et de la cochlée dans une tête humaine.
Lindsay P., & Norman D., 1980, figure 4.1.

Le *système vestibulaire* comprend trois canaux semi-circulaires grossièrement positionnés dans les trois plans de l'espace (figure 3.2) et munis de capteurs sensibles aux accélérations de la tête. Il comprend également l'utricule et le saccule, dont les capteurs, sensibles à la pesanteur, positionnés dans des plans à peu près perpendiculaires entre eux, nous renseignent sur la position du corps dans l'espace. Comme les enveloppes de ces capteurs baignent dans le même liquide, *la périlymphe*, que le canal cochléaire de la cochlée, on voit que les organes de l'équilibration ont un lien privilégié avec ceux de l'audition.

La *cochlée* ou limaçon, présente l'allure d'un escargot dont la coquille aurait deux tours et demi.

La section transversale (figure 3.7) fait apparaître trois parties : la rampe vestibulaire, le canal cochléaire et la rampe tympanique, qui, du point de vue fonctionnel, correspondent à deux circuits hydrodynamiques distincts[3].

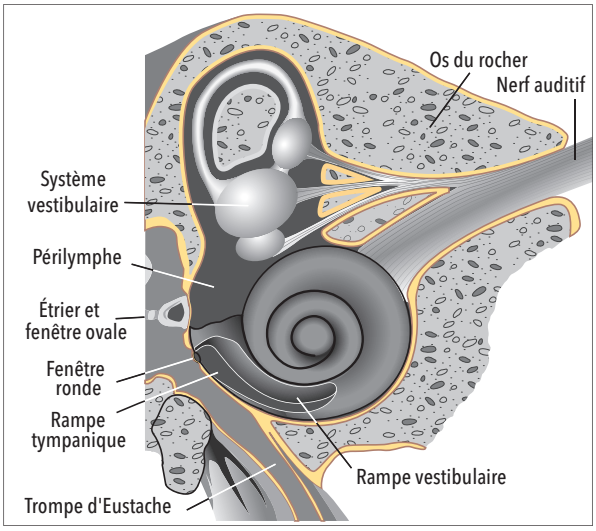


Figure 3.6 Coupe schématique de l'oreille interne montrant le système vestibulaire et la cochlée logés dans l'os du rocher, baignant dans la périlymphe.

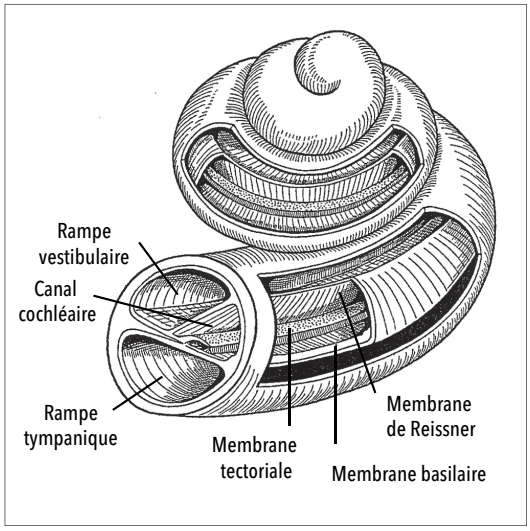


Figure 3.7 Dessin de la cochlée avec les deux rampes vestibulaire et tympanique, et le canal cochléaire.

D'après Legent, F., Perlemuter, L., & Vandenbrouck, C., 1968, « Labyrinthe membraneux », planche H.

Le premier, empli de périlymphe, est constitué de la rampe vestibulaire qui communique au sommet par un trou (l'hélicotrème) avec la rampe tympanique (figure 3.8). Les parois externes de l'organe étant enchâssées dans l'os du rocher, et le liquide étant incompressible, la transmission des vibrations ne peut se produire que parce qu'il existe deux opercules membraneux : la *fenêtre ovale* à laquelle aboutit l'étrier et la *fenêtre ronde* située à la partie terminale de la rampe tympanique (figure 3.8).

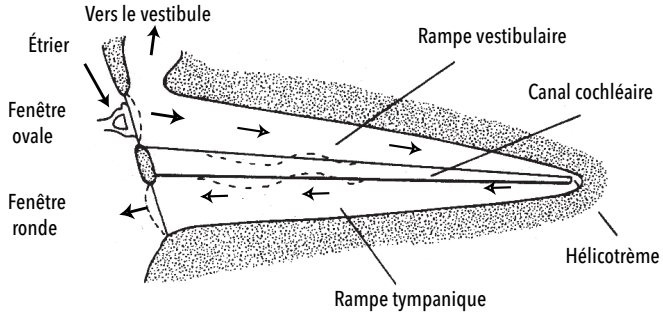


Figure 3.8 Schéma fonctionnel de la cochlée.

Pickles, J., 1982, figure 3.1-C.

3. Pour accéder à une bonne représentation du système auditif et de son comportement dynamique, nous recommandons le site Internet de l'Inserm, www.cochlea.org, qui propose des animations selon deux niveaux de présentation : grand public et professionnels.

Entre ces deux rampes se situe le canal cochléaire, cœur de l'organe, qui est un conduit étanche empli d'endolymphe. Le canal cochléaire possède deux parois membraneuses déformables : la membrane basilaire et la membrane de Reissner. À l'intérieur du canal cochléaire se trouvent les cellules nerveuses sensibles aux mouvements vibratoires produits par les déplacements de la périlymphe.

L'étude détaillée du fonctionnement de la cochlée dépasserait le cadre de cet ouvrage. Deux éléments vont toutefois retenir notre attention : la membrane basilaire et les cellules ciliées.

La membrane basilaire. Elle est constituée de fibres radiales dont la longueur augmente depuis la base jusqu'au sommet de la cochlée (figure 3.9). En réponse à une excitation sinusoïdale, de basse fréquence, l'enveloppe du maximum de déplacement de la membrane se situe vers le sommet, zone où celle est large et souple. Au contraire, lorsqu'il s'agit de hautes fréquences, c'est la portion près de la base, zone plus étroite et plus raide, qui répond à l'excitation.

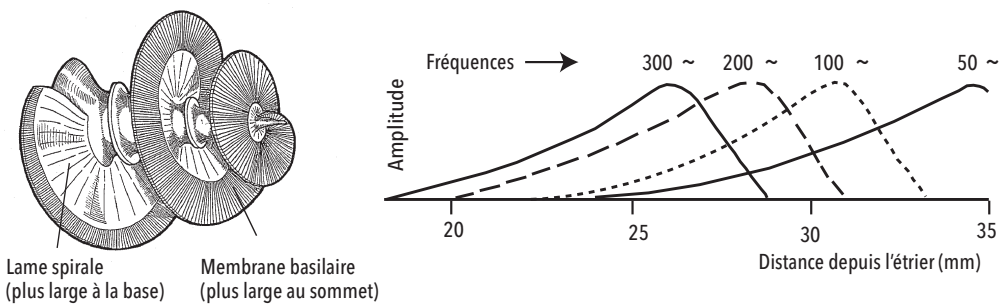


Figure 3.9 La membrane basilaire – À gauche, la lame spirale (os), plus large à la base et la membrane basilaire, plus large au sommet. À droite, réponse de la membrane basilaire à une excitation sinusoïdale. Pour un son pur, la position du maximum de déplacement, donc la configuration spatiale, correspond simultanément à la fréquence et au maximum d'énergie dans le spectre.

À gauche : Legent & al., 1968, « Labyrinthe osseux », Planche L

À droite : Pickles, 1982, figure 3.8, (selon Békésy, 1960).

À une excitation sonore sinusoïdale donnée correspond donc une configuration vibratoire spécifique de la membrane basilaire, que l'on qualifie de codage spatial ou tonotopique (figure 3.9, droite). Ce type de codage associe la position du maximum de déplacement (donc d'énergie vibratoire) dans la cochlée à la fréquence de vibration. Nous verrons plus loin (chapitre 6, § 1.3) comment cette observation peut s'interpréter dans l'analyse des sons complexes, et en particulier au sujet du lien ambigu entre sensation de timbre et sensation de hauteur.

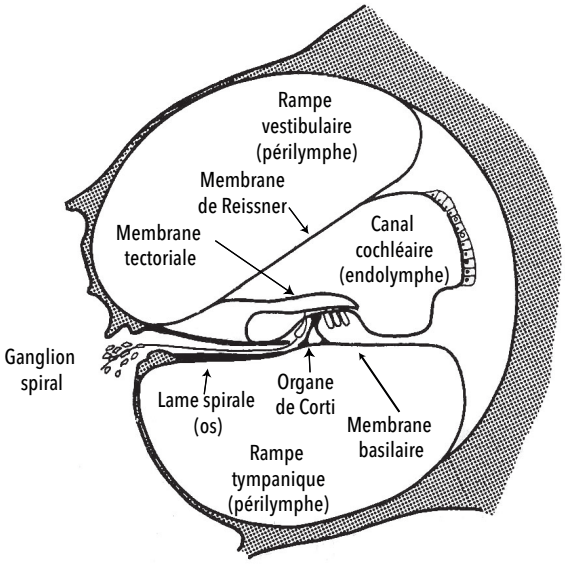


Figure 3.10 Coupe schématique de la cochlée. C'est dans le canal cochléaire, situé entre les rampes vestibulaire et tympanique, que se trouvent les cellules nerveuses ciliées dont l'excitation produit les signaux véhiculés par le nerf auditif.

D'après Pickles, J., 1982, figure 3.31-B.

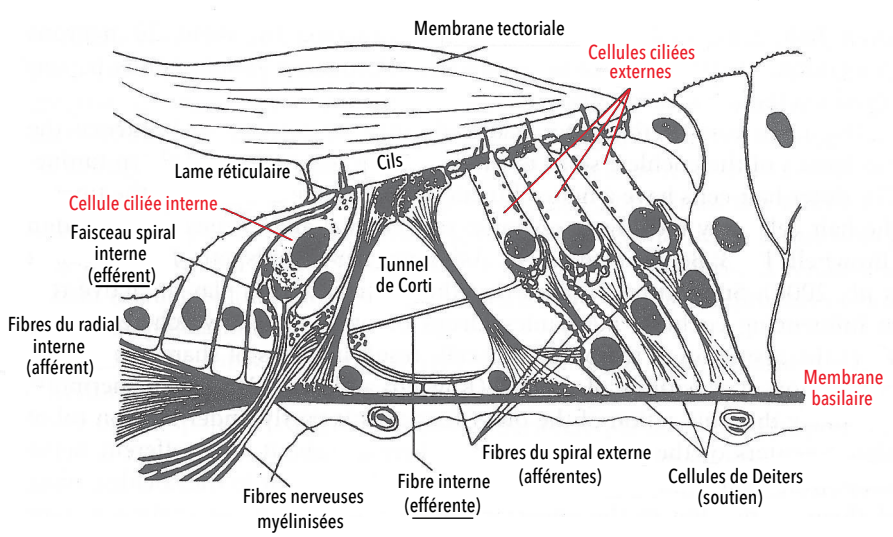


Figure 3.11 Détail de l'organe de Corti.

D'après Moore, B., 2013, figure 1.14.

Les cellules ciliées. Le canal cochléaire renferme deux groupes de cellules sensorielles disposées de part et d'autre d'un espace appelé tunnel de Corti (figures 3.10 et 3.11).

Les *cellules ciliées internes* (au nombre d'environ 3500, possédant chacune quelque 40 stéréocils) sont directement connectées aux fibres du nerf auditif. Elles assurent principalement la transduction des mouvements mécaniques en impulsions nerveuses à destination des centres supérieurs.

Les *cellules ciliées externes*, trois fois plus nombreuses (environ 12000, possédant chacune quelque 140 cils), sont sous le contrôle des centres supérieurs. Elles se contractent sous l'effet des fibres efférentes (soulignées sur la figure 3.11) et peuvent modifier les caractéristiques mécaniques de la cochlée[4]. Ainsi donc, dès les premières étapes de sa transmission, la réponse sensorielle est sous le contrôle d'informations descendantes provenant des centres supérieurs.

La distribution des zones de la membrane basilaire sensibles aux différentes fréquences se retrouve dans les fibres du nerf auditif. On a pu vérifier que les fibres isolées ne répondent qu'à certaines fréquences caractéristiques.

1.4. Les voies nerveuses : de la cochlée au cortex auditif

Les impulsions nerveuses issues du nerf auditif transitent dans un réseau complexe de structures nerveuses (noyau de l'olive supérieure, noyau cochléaire, colliculus inférieur et corps genouillé médian) au sein desquelles les signaux subissent des traitements spécifiques : détection du début et de la fin d'un son, corrélations temporelles et topologiques et, plus particulièrement, des comparaisons entre les signaux provenant de l'oreille droite et ceux de l'oreille gauche dont les voies se croisent à plusieurs reprises (figure 3.12).

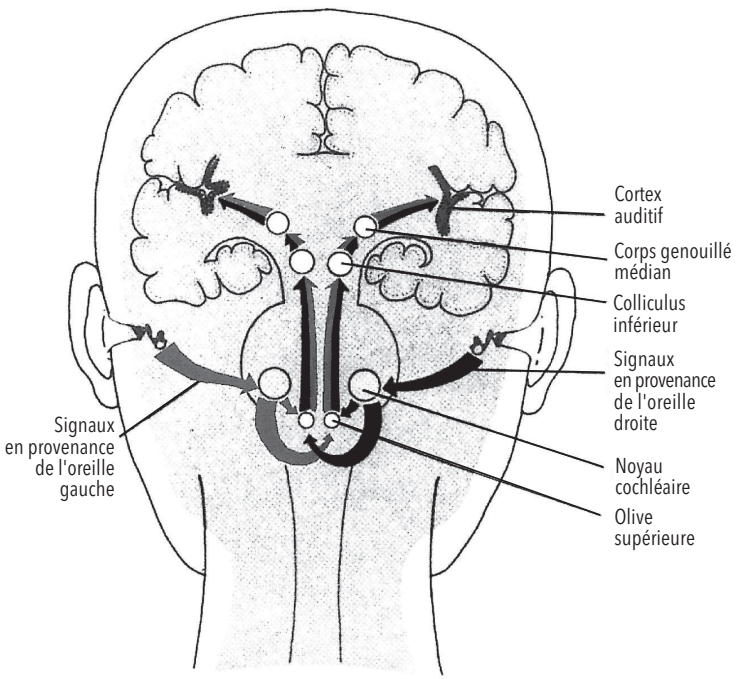


Figure 3.12 Voies nerveuses, de l'oreille interne au cortex auditif.
Lindsay P., Norman, D., 1980, figure 6.39 a

4. On explique ainsi le fait que la sélectivité fréquentielle est très supérieure chez l'animal vivant.

À l'étage supérieur, celui des hémisphères cérébraux, sont dévolus les traitements dits *cognitifs*, mettant en jeu l'ensemble des opérations mentales qui caractérisent un individu humain. L'écoute est une fonction active qui implique, dans l'intervalle de temps où elle se déroule, la conscience d'un événement acoustique que l'on peut reconnaître – parce qu'on en a fait l'expérience et qu'il a été mémorisé – et dont on peut anticiper le devenir, en adaptant l'oreille et les prétraitements pour sélectionner certains éléments du signal qui nous intéressent.

S'il est évident que ces opérations ont pour substrat des impulsions nerveuses et des circuits neurophysiologiques, les chercheurs sont encore loin de fournir des données ou des modèles explicatifs, et les discussions concernant les localisations éventuelles de certains traitements sont âpres. Toutefois, les observations consécutives aux accidents cérébraux semblent confirmer que les traitements de haut niveau, tels que la discrimination entre sons périodiques et bruits, mélodie et rythme et les diverses interprétations de la parole, concernent des zones corticales différenciées, en particulier chez les musiciens professionnels.

Les observations effectuées avec les techniques de magnétoencéphalographie (MEG) et de tomographie par émission de positons (PET), et surtout l'imagerie par résonance magnétique fonctionnelle (fIRM) ont fourni des résultats spectaculaires depuis une dizaine d'années. Mais l'interprétation des résultats diffère selon les auteurs et surtout selon les protocoles expérimentaux et les types de signaux utilisés. Pour certains (Perry *et al.*, 1999 ; Zatorre *et al.*, 2002), les aspects temporels, séquentiels, analytiques concerneraient plutôt l'hémisphère gauche (pour un droitier) et les aspects spectraux, synthétiques, hédoniques plutôt l'hémisphère droit. D'autres (Maidhof et Koelsch, 2011 ; Schön *et al.*, 2005), qui étudient les traitements croisés de la parole et de la musique, ne constatent pas de préférence hémisphérique pour l'un ou l'autre de ces signaux, mais des implications différentes selon le type de tâche demandée et surtout selon le degré d'attention induit par la consigne.

> *Finally, the extent to which neural representations of auditory streams in (and below) the auditory cortex are influenced by selective attention deserves further investigation.The abundance of descending (efferent) connections in the auditory system provides ample opportunity for "top-down" influences, and makes it quite possible that effects of selective attention affect early stages of the neural analysis of auditory scenes.*[5]
>
> Shamma, S., Micheyl, C., 2011, p. 365.

Il faut mentionner aussi l'approche proposée depuis une quinzaine d'années par plusieurs auteurs (voir Chi, T., Ru, P., & Shamma, A., 2005), qui ont développé une modélisation des traitements du signal sonore fondée sur les observations neurophysiologiques. Une des originalités du postulat consiste à poser que les unités cellulaires effectuent en parallèle l'analyse bidimensionnelle – spectrale et temporelle – du signal sonore, et que leurs caractéristiques de résolution en temps et en fréquence diffèrent, à un instant donné. Cette analyse « multirésolution », se développant selon deux étapes – de la cochlée au thalamus pour le spectrogramme auditif et du thalamus au cortex pour l'analyse spectrotemporelle proprement dite –, rend compte d'une bonne discrimination des signaux complexes (sons musicaux et parole). De surcroît, les observations en temps réel sur l'animal (le furet) montrent que les réponses corticales varient selon l'attente (espérée ou redoutée) du

5. « Finalement, l'importance avec laquelle les représentations neurales des flux auditifs produites dans le cortex auditif (et aux étages inférieurs) sont influencées par l'attention sélective mérite de plus amples recherches. L'abondance des connexions descendantes (efférentes) dans le système auditif témoigne des nombreuses possibilités d'influences « top-down » et plaide en faveur d'une influence de l'attention sélective aux stades précoces de l'analyse neurale des scènes auditives. » Traduction M. C.

signal, ce qui confirme le rôle des voies nerveuses descendantes (efférentes) dans le traitement des signaux sonores.

La figure 3.13 récapitule de façon schématique les principales fonctions des différents étages de la transmission dans le système auditif.

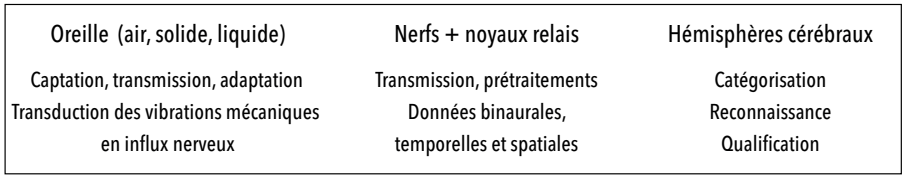


Figure 3.13 Récapitulation schématique des étapes du traitement des signaux sonores dans le système auditif. Dans la réalité, de nombreuses connexions existent entre les trois blocs arbitrairement séparés dans la figure.

1.5. Le codage de l'information sensorielle

1.5.1. La transduction : réponse nerveuse à une excitation sensorielle

L'étape clé est celle de la conversion de l'information vibratoire issue du monde extérieur en influx nerveux. Deux acteurs sont essentiels : la cellule sensorielle, qui capte le phénomène vibratoire, et la cellule nerveuse, qui communique avec la précédente au niveau d'une zone de contact appelée synapse. À l'issue de la conversion, les signaux transitant dans le système nerveux sont tous de nature binaire (voir encadré page 92). L'information sensorielle est donc transmise sous une forme codée qu'il s'agit de décrypter.

En ce qui concerne l'audition, on considère le codage *spatial* et le codage *temporel*, ainsi qu'une grande variété de représentations de l'excitation selon les regroupements des réponses et leurs synchronisations.

1.5.2. Le codage spatial ou tonotopique de la fréquence et de l'intensité

Fréquence. Selon leur fréquence, les vibrations sonores affectent différentes parties de la cochlée. Un son de 300 Hz excite les cellules ciliées proches du sommet de l'hélice, alors qu'un son de 4000 Hz excite les cellules ciliées proches de la base.

La sélectivité de la réponse dépend de l'action des cellules ciliées externes (voir figure 3.11, page 88), mais, en pratique, l'excitation d'un son pur affecte plusieurs cellules, donc plusieurs fibres nerveuses. Cependant, pour une fréquence donnée, on peut déterminer la fibre dans laquelle la réponse est maximale, ou encore la fréquence caractéristique de cette fibre. Cette observation, jointe à d'autres, confirme que l'information topologique du site de la membrane basilaire est conservée dans le nerf auditif et se retrouve aux différentes étapes de la chaîne de transmission : c'est le codage tonotopique de la fréquence.

Intensité. L'augmentation d'intensité du signal provoque un élargissement de la zone d'excitation de la membrane basilaire : le *pattern d'excitation* des neurones activés change selon l'intensité. Mais l'élargissement ne croît pas de façon uniforme. Lorsque l'intensité augmente, la zone impliquée s'étend plus largement vers la base de la membrane basilaire, c'est-à-dire vers les hautes fréquences. L'asymétrie de l'élargissement se retrouve dans le nombre des fibres répondant à

des excitations d'intensité croissante, ainsi que dans la forme des *patterns de masquage*. Elle est en accord avec le fait que les sons aigus sont plus facilement masqués que les sons graves (voir encadré page 102).

Pour interpréter ces expériences, on représente la membrane basilaire comme étant constituée d'un ensemble de filtres de large bande appelés *bandes critiques* (voir § 2.2, page 98).

On connaît peu de choses sur le comportement de la membrane basilaire en réaction aux sons complexes et aux sons instationnaires. La réponse à un clic serait une onde progressive dont la forme change au cours de son déplacement. À ces indications, il faut ajouter l'action de cellules inhibitrices et/ou facilitatrices qui, vraisemblablement, peuvent également modifier les caractéristiques de base, sous l'action de commandes provenant du cerveau (voies efférentes). Ces mécanismes peuvent être invoqués pour expliquer l'abaissement du seuil de perception d'un son, lorsque celui-ci est prévisible.

Réponse d'une cellule sensorielle[1]

Une cellule sensorielle est une sorte de centrale chimique. Lors d'une excitation sonore, la flexion des cils terminaux d'une cellule de Corti provoque une dépolarisation de la membrane par échange ionique. Les ions potassium sortant sont remplacés par des ions sodium. Il en résulte une différence de potentiel au niveau de la synapse, zone de contact entre la cellule ciliée et les terminaisons d'un neurone du nerf auditif. Cette différence de potentiel est à l'origine d'une impulsion, ou potentiel d'action, qui se propage dans l'axone du nerf. Rapidement, le mouvement des ions s'inverse et la cellule retrouve son état initial, prête à être excitée de nouveau.

Pour qu'une cellule réponde, deux conditions sont nécessaires :

1/ le niveau de l'excitation doit être suffisamment élevé pour déclencher le processus. En deçà d'un seuil minimal ou seuil absolu, la cellule ne répond pas ;

2/ la cellule ne peut répondre à une deuxième stimulation qu'après une certaine durée de récupération, appelée période réfractaire. Celle-ci varie de 0,5 à 1 ms selon les récepteurs et selon le type d'excitation.

Quelle que soit la cellule considérée, la nature de la réponse est toujours binaire : il se produit une impulsion ou non. La cadence des impulsions augmente avec la croissance d'intensité du stimulus. Dans la réalité, une excitation affecte toujours un groupe plus ou moins important de cellules dont le comportement global devient rapidement complexe. Des phénomènes d'inhibition – certaines cellules bloquent leurs voisines – de facilitation, de synchronisation de groupes de cellules, permettent d'envisager une grande variété d'encodage des différenciations qualitatives de l'excitation.[1]

1. Buser & Lestienne, 2001, p. 31.

1.5.3. Les codages temporels de la fréquence et de l'intensité

La deuxième modalité de codage, temporelle, concerne la distribution des décharges nerveuses dans les axones.

Intensité et durée des signaux. Nous avons déjà fait mention du codage temporel, le plus simple et le plus répandu dans les organismes vivants. Dès qu'un signal parvient au récepteur, les cellules produisent des décharges dont le nombre augmente avec l'intensité et qui cessent avec l'arrêt du signal. Il existe aussi des cellules qui ne répondent qu'au début du signal (*on*) ou à la fin (*off*). On décrit également des types de cellules ayant des seuils d'excitabilité différents, ce qui permet de rendre compte de la sensibilité aux variations d'intensité et d'imaginer comment est représentée la gamme dynamique des réponses.

Fréquence. Les sons périodiques possèdent une structure temporelle remarquable qui se répète à intervalle de temps constant : la période. Plusieurs auteurs ont montré que la structure temporelle des influx nerveux transitant dans les fibres du nerf auditif se synchronise sur la période T et sur ses multiples 2T, 3T, etc. (figure 3.14, partie supérieure). Le relevé statistique de la distribution des intervalles de temps entre les impulsions transitant dans une fibre, en réponse à l'excitation par un son sinusoïdal, présente un maximum pour la valeur de la période du signal d'excitation (figure 3.14, partie inférieure).

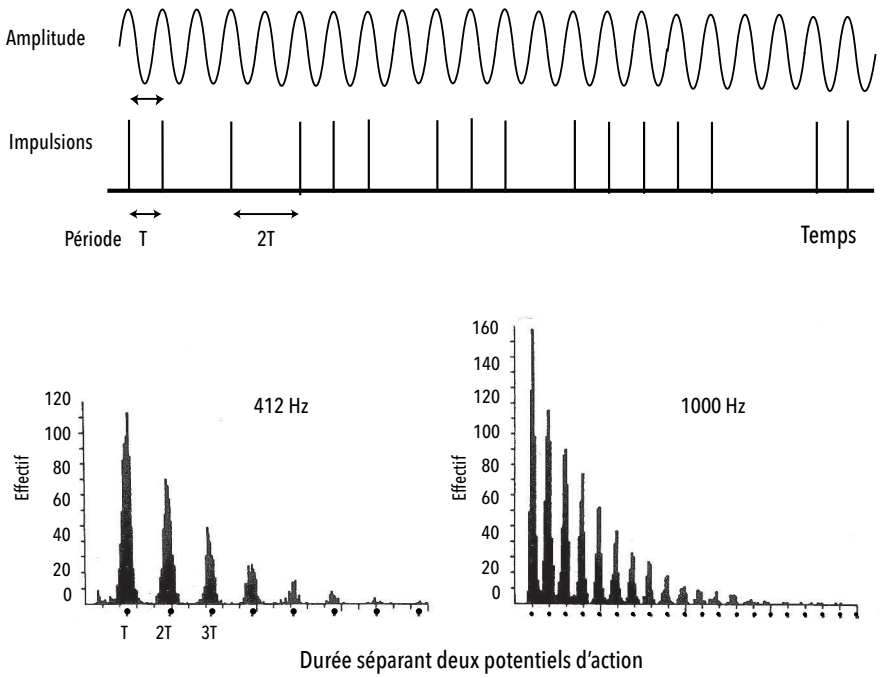


Figure 3.14 En haut, schéma d'une forme d'onde sinusoïdale de période T et, en dessous, les impulsions correspondantes dans une fibre nerveuse, déclenchées par les maxima d'amplitude. Toutes les occurrences ne sont pas détectées. En bas, histogrammes expérimentaux de la réponse d'une fibre nerveuse à l'excitation de deux sons purs : 412 Hz et 1000 Hz. En abscisse : intervalle de temps séparant deux décharges. En ordonnée : nombre de réponses cumulées au cours de 10 enregistrements de chacun 1 s. La répartition statistique des décharges confirme l'organisation temporelle des réponses : le plus grand nombre est calé sur la période, ensuite on trouve les multiples.

En bas : d'après Demany, L., 1989, p. 49 (selon Rose & coll., 1967).

Ce mode de codage est nommé codage temporel de la fréquence, par opposition au codage spatial ou tonotopique. Il fournit une information extrêmement précise, mais il est limité vers les hautes fréquences : entre 4 et 5 kHz, la synchronisation se dégrade puis disparaît (voir page 112).

Mentionnons enfin le codage par latence temporelle, dont l'intérêt apparaît lors de la comparaison d'influx nerveux ayant parcouru des circuits différents (Buser & Lestienne, 2001).

La présentation simplifiée que nous venons de faire montre déjà que plusieurs modalités de codage existent pour un paramètre physique donné tel que l'intensité ou la fréquence.

L'intensité peut être représentée par la quantité de décharge des influx nerveux d'une cellule, mais aussi par l'augmentation du nombre de cellules excitées.

La fréquence peut être représentée par la réponse de cellules situées à des emplacements spécifiques de la membrane basilaire, mais aussi par l'intervalle temporel entre les influx nerveux de certains neurones.

L'étude des traitements aux différentes étapes de la transmission nerveuse rend compte de différenciations plus fines, dues à des regroupements sur la base d'organisations spatiales et temporelles dans les noyaux relais.

Il semble que les données temporelles y jouent un rôle prépondérant (voir les ouvrages de Buser, 2001 ; Moore, 2013 ; les articles de Patterson, page Web personnelle).

1.5.4. Les non-linéarités de l'oreille

Un grand nombre d'observations révèlent un fait important : les organes de saisie et de traitement des signaux ont un comportement non linéaire particulièrement accusé lorsque l'amplitude d'excitation est grande (voir Glossaire).

Au niveau de l'oreille moyenne, les vibrations du tympan et de la chaîne des osselets peuvent provoquer des distorsions harmoniques et des sons de combinaison (en particulier les sons différentiels : voir chapitre 8, page 415).

Dans l'oreille interne, qu'il s'agisse du comportement de la membrane basilaire ou des décharges dans les fibres nerveuses, plusieurs expériences ont montré que les réponses ne sont pas proportionnelles aux variations d'amplitude des stimuli (voir Moore page 29, figure 1.12 pour les sons sinusoïdaux et page 32, figure 1.13 pour les clics).

En bref, et ce point est important, la réponse à des excitations complexes n'est pas simplement la somme des réponses obtenues pour des stimuli élémentaires.

2. Les caractéristiques psychophysiques de l'oreille humaine

Comme tout organe sensoriel, l'oreille ne réagit aux vibrations sonores que dans une certaine étendue de fréquences et d'énergie vibratoire, et elle est limitée par des *seuils absolus*. À l'intérieur de la zone de réception, la discrimination plus ou moins fine des sensations sonores dépend des *seuils différentiels*.

2.1. Les seuils absolus

2.1.1. L'aire audible en intensité et en fréquence

L'aire audible est, par définition, la zone de fréquence à l'intérieur de laquelle l'oreille peut capter des sons d'intensité et de fréquence variables (figure 3.15). La dynamique totale de l'oreille s'étend entre le *seuil d'audibilité* et le *seuil de douleur.*

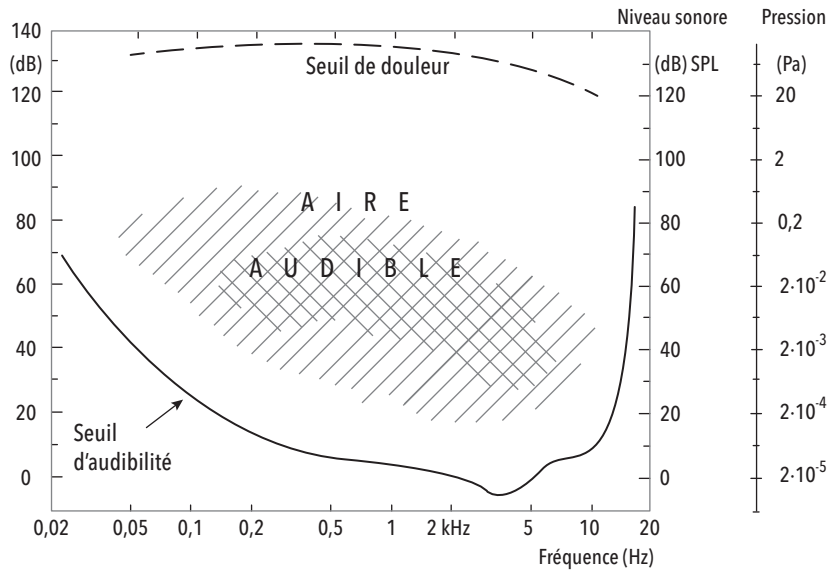


Figure 3.15 L'aire audible est la zone des fréquences, du grave à l'aigu, comprises entre le seuil d'audibilité et le seuil de douleur. Le maximum de sensibilité se situe aux alentours de 3000 Hz. Pour un son d'intensité donnée, constante, la sensation d'intensité perçue varie avec la fréquence.

D'après Leipp, E., archives LAM.

Le *seuil d'audibilité* est, pour un son pur, le minimum de pression déclenchant une sensation sonore. Sa valeur varie selon les fréquences. Le seuil d'audibilité minimum se trouve aux alentours de 3000 Hz, zone de fréquences qui est statistiquement la zone de plus grande sensibilité de l'oreille. La valeur 0 des échelles de mesure a été définie pour la fréquence 1000 Hz[6] et correspond à des variations de pression de 2×10^{-5} Pa, soit une amplitude vibratoire de 10 picomètres, dimension comparable au diamètre d'une molécule d'hydrogène. Le tracé du seuil de sensibilité varie selon les méthodes de mesure (écoute en champ libre, au casque ou avec écouteurs internes).

6. En conséquence, le seuil absolu pour 3000 Hz a une valeur négative.

Le *seuil de douleur* apparaît quand la pression sonore atteint 20 Pa. Il diffère peu selon les fréquences.

Par référence à l'aire d'audibilité humaine, on définit les ultrasons de fréquence très élevée (non perçus par l'oreille humaine) et les infrasons, ressentis comme vibrations, mais trop graves pour procurer une sensation sonore proprement dite.

2.1.2. Les tests d'oreille : les audiogrammes

Plusieurs méthodes permettent d'évaluer les capacités auditives d'un sujet. Le test le plus courant est *l'audiogramme tonal,* que l'on effectue pour chaque oreille avec un équipement normalisé (audiomètre et casque étalonné). L'appareil délivre une suite de sons purs d'intensité variable. Pour chacun d'eux, le praticien recherche le seuil minimum d'intensité perçue par un sujet et le compare ensuite à un seuil de référence, défini selon une norme internationale : c'est l'axe du 0 dB sur la figure 3.16.

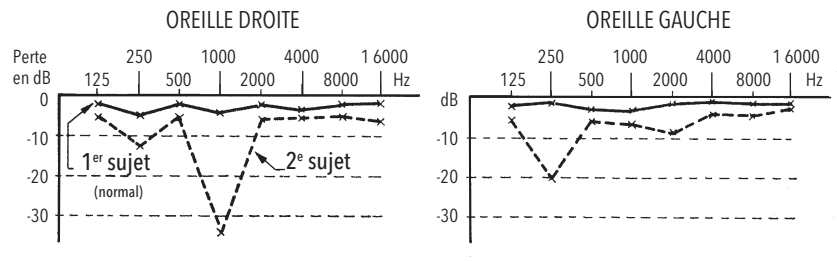


Figure 3.16 L'audiogramme tonal permet d'évaluer la sensibilité de chaque oreille avec des sons sinusoïdaux calibrés. On détecte ainsi les « trous » de l'audition. Cet audiogramme est très insuffisant ; on le complète par d'autres tests.

Leipp, E., 2010, figure 51.

Chez les sujets ayant une audition normale, des écarts d'environ 10 dB de part et d'autre de cette référence s'observent couramment. La courbe est rarement plate. Pour certaines personnes elle présente une pente ascendante vers les hautes fréquences ; pour d'autres elle est descendante vers les basses fréquences. Pour un même individu, les deux oreilles ont généralement des courbes très similaires.

L'audiogramme tonal est un test assez fruste avec lequel on n'explore qu'un nombre très limité de fréquences. Il permet de vérifier qu'un sujet entend normalement, mais ne saurait fournir aucune indication sur ses capacités musicales. D'autres tests existent, comme *l'audiogramme verbal* (dictée de listes de mots normalisés), qui donne une appréciation des pertes auditives pendant la conversation, ou encore *l'audiogramme par transmission osseuse,* qui court-circuite la transmission de l'oreille moyenne et permet d'évaluer directement l'état de l'oreille interne.

Pour prospecter la capacité à discriminer des sons successifs, on peut compléter l'audiogramme tonal par un test temporel (voir page 125).

2.1.3. Les pertes de la sensibilité auditive : fatigue, vieillissement, surdités

L'élévation du seuil de sensibilité auditive avec l'âge est inéluctable et affecte prioritairement les hautes fréquences. Dès l'âge de 25 ans, la sensibilité aux fréquences supérieures à 16 kHz s'émousse. Ensuite, la perte atteint 14, puis 10 kHz

(figure 3.17). Ces modifications ne sont pas perçues par le sujet, qui s'adapte en permanence[7]. Elles n'ont qu'un faible retentissement sur les performances musicales, mais peuvent affecter les jugements sur la sonorité des instruments de musique riches en harmoniques, comme le violon, le clavecin, ou encore la voix chantée.

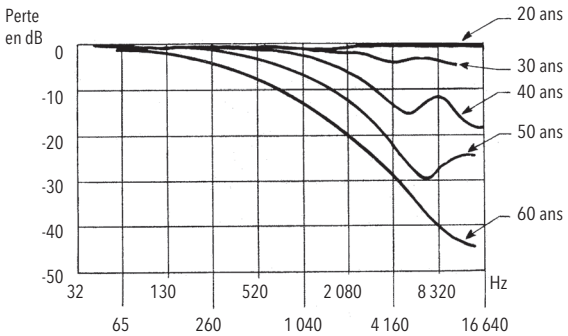


Figure 3.17 Perte d'audition avec l'âge. Statistiquement, l'oreille humaine perd graduellement sa sensibilité dans l'aigu avec l'âge : c'est ce que traduit ce diagramme. Les pertes varient considérablement d'un individu à l'autre.

Leipp, E., 2010, figure 52.

Lorsque la conversation devient difficile, c'est le signe d'une élévation du seuil d'audibilité qui affecte la région inférieure à 4 kHz, et la presbyacousie devient une gêne sociale. Simultanément la dynamique générale de l'oreille se réduit : on constate qu'un sujet âgé est aussi plus sensible aux bruits intenses.

Une perte temporaire de la sensibilité auditive est observable après l'écoute prolongée de sons intenses. Chez un sujet jeune, la fatigue induite par une exposition occasionnelle disparaît après le repos d'une bonne nuit. Mais la pratique régulière d'une musique à forte intensité, l'habitude inconséquente d'écouter au casque à fort niveau produisent des fatigues à répétition qui provoquent un nombre grandissant de surdités partielles, malheureusement irréversibles.

L'écoute de simulations de surdités est instructive. Les Sons **3.1** et **3.2** font entendre une voix lisant un texte avec différents filtrages. L'interprétation de ces filtrages est très complexe, puisque la compréhension de la parole fait intervenir des traitements de haut niveau, et en particulier la suppléance par le contexte général de la phrase. Ces exemples permettent d'entendre les changements de qualité de la voix selon les types de filtrages.

Son 3.1 (52″)

Filtrage passe-bas

Son 3.2 (58″)

Filtrage passe-haut

Il existe différents types de surdités. Lorsque seules les parties externe et moyenne de l'oreille sont atteintes : inflammation ou détérioration du tympan, blocage de la chaîne des osselets de l'oreille moyenne (otospongiose), la restauration chirurgicale est très satisfaisante. Lorsque les cellules ciliées de l'oreille interne sont atteintes, la chirurgie mécanique est impuissante. Depuis une vingtaine d'années s'est développée une technique qui consiste à remplacer les cellules sensorielles inactives par des implants cochléaires. Elle nécessite un long et lourd apprentissage pour recréer un nouveau codage des sensations sonores. Il existe également des surdités dites centrales qui correspondent à l'atteinte des centres nerveux de traitement de l'information sonore. Le sujet ne reconnaît plus la parole ou la musique alors que son oreille est saine. Ce sont toujours des cas complexes, parfois réversibles, ou susceptibles de suppléances, d'autres zones du cerveau prenant le relais pour remplacer fonctionnellement les zones lésées.

7. À l'époque des téléviseurs à tube cathodique, la fréquence de balayage de 15 625 Hz en France (625 × 25) était un test impitoyable.

2.2. L'oreille considérée comme un banc de filtres : la notion de bandes critiques

2.2.1. À propos du filtrage et de la sélectivité en fréquence

Remarque préliminaire. Une opération de filtrage modifie la répartition de l'intensité sur les différentes fréquences. Le temps est également affecté puisqu'on définit la réponse temporelle d'un filtre. Selon les cas le filtrage affectera plus spécifiquement l'un ou l'autre de ces paramètres.

Parmi les capacités auditives, celle qui permet de différencier les sons selon leur hauteur intrigue depuis longtemps les chercheurs en psychophysique. C'est en partie grâce aux différences de la hauteur fondamentale que nous sommes capables de distinguer rapidement une voix d'homme d'une voix de femme ou d'enfant, ou encore d'interpréter certaines intentions de l'interlocuteur à la seule écoute de son intonation. Dans la musique européenne, les différenciations des hauteurs jouent un rôle majeur ; elles peuvent être extrêmement fines – de l'ordre du 1/300 d'octave – et, pour certains sujets, peuvent même donner lieu à une mémorisation de la fréquence (oreille dite absolue).

Un postulat initial de la psychophysique est que la fréquence d'un signal conditionne la sensation de hauteur. Les expériences sont faites avec des sons purs, et il s'agit de rechercher dans le système auditif des analyseurs de fréquence pouvant expliquer nos capacités de discrimination des hauteurs des sons.

Les difficultés surgissent lorsqu'il faut rendre compte de la perception de sons réels complexes. D'une part, plusieurs des éléments de la chaîne auditive ayant un comportement non linéaire, on sait que la perception d'un son complexe ne peut pas être représentée par la somme des effets dûs à ses composantes isolées. D'autre part, bon nombre de sons réels communs donnent une sensation de hauteur non ambiguë bien que n'ayant pas (ou peu) d'énergie à la valeur de la fréquence fondamentale, enfin d'autres sont inharmoniques.

Cette section présente un bilan des connaissances récentes et propose une discussion sur la perception des sons complexes. À la suite de Helmholtz, qui, le premier, a émis l'hypothèse que l'oreille effectuait une analyse des fréquences, Békésy (1960) a attribué l'essentiel de la fonction de *filtrage en fréquence* à la structure particulière de la membrane basilaire. Sans être fondamentalement remise en cause, cette fonction apparaît aujourd'hui sous un jour plus complexe. Sa mise en œuvre semble impliquer la coopération de l'ensemble des différentes parties du système auditif, depuis l'oreille interne jusqu'aux centres supérieurs qui interagissent, comme on l'a vu (voir figures 3.1 et 3.11), via les voies efférentes. Sur la base des observations physiologiques et au vu des résultats de nombreuses expériences de psychoacoustique, différents concepts ont été élaborés. L'emploi du terme « concept » indique ici que l'on peut objectiver les résultats de traitements opérés à différents niveaux de la chaîne de transmission, sans pour autant attribuer avec certitude des rôles définis à des structures spécifiques.

Parmi les notions qui permettent de comprendre le traitement des signaux par l'oreille nous retiendrons celle de *bande critique.*

2.2.2. Le traitement dual de l'intensité et de la fréquence : les bandes critiques

La notion de bande critique proposée par Fletcher en 1940 a été confirmée depuis par de nombreuses expériences. Elle permet d'expliquer les réponses des auditeurs lors d'expériences d'évaluation de la sensation d'intensité de sons complexes, ou lors d'expériences de masquage.

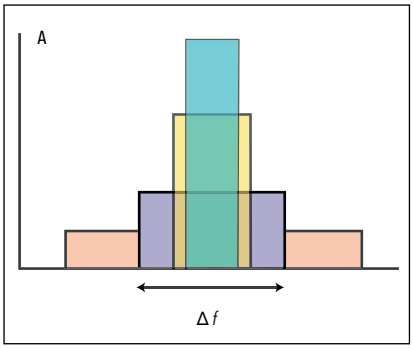


Figure 3.18 Un bruit de bande étroite de largeur variable, et d'énergie globale constante.

Une première expérience (Son **3.3***)[8] consiste à comparer le niveau sonore de deux bruits. Le premier bruit, de largeur Δf_1 est fixe ; il sert de référence. La largeur de bande Δf_2 du second bruit croît, mais son amplitude est rajustée de sorte que l'énergie globale reste constante (figure 3.18). Pendant les premières présentations, les deux bruits sont perçus à intensité sonore équivalente. Pour une certaine valeur de Δf_2 le deuxième bruit paraît plus fort, puis son intensité continue de croître avec l'augmentation de Δf.

Son 3.3* (25")

Pour le détail de l'expérience, voir légende du son (page 135)

Une autre expérience porte sur le masquage simultané. Un son pur de fréquence f_1 et un bruit dont la largeur de bande Δf est centrée sur f_1 sont émis simultanément. Le niveau du bruit est ajusté de sorte que le son soit tout juste perceptible : c'est la valeur du seuil de détection du son pur. Ensuite Δf croît, le niveau du bruit est modifié pour conserver une densité spectrale constante et il faut accroître celui du son pur pour qu'il reste perceptible. On répète l'expérience jusqu'au moment où l'on constate que le seuil de détection du son reste constant, bien que la largeur de bande du bruit continue de croître.

On en conclut que la modalité d'intégration de l'intensité sonore dépend d'une certaine largeur en fréquence appelée *bande critique*. En d'autres termes, les vibrations sonores interagissent différemment lorsqu'elles affectent des sites de la membrane basilaire voisins ou distants. Pour les deux expériences la largeur de la bande Δf pour laquelle s'opère la transition définit celle de la bande critique pour la fréquence centrale considérée.

2.2.3. La largeur des bandes critiques

La détermination de la largeur en fréquence des bandes critiques a fait l'objet d'un très grand nombre d'expériences. Zwicker (1999) a défini 24 bandes (*Barks*), ayant chacune une position définie dans le spectre. D'autres (Moore, 2013) considèrent plutôt qu'il s'agit de zones mobiles s'étendant de part et d'autre de la fréquence centrale, les bandes voisines se chevauchant partiellement. La figure 3.19 (page suivante) montre les relations qui lient la fréquence d'excitation et la largeur (en fréquence) des bandes critiques.

La droite en trait plein montre que la relation est constante pour les fréquences supérieures à 1000 Hz, mais augmente vers les sons de plus basse fréquence. À l'examen du diagramme, on lit pour les valeurs 0,5, 1 et 2 kHz de la fréquence centrale les largeurs de bande suivantes : Δf = 80 Hz, 130 Hz et 250 Hz. Exprimons ces trois intervalles en pourcentage de la fréquence centrale. On trouve : 16 %[9],

8. Cet exemple, et tous ceux qui sont marqués d'un astérisque, sont à écouter au casque. À noter : la fatigue infligée par ces stimuli.
9. En effet, le rapport 80/500 = 0,16 soit 16 %.

13 % et 12 %. Sachant que 12 % correspond à un intervalle d'un ton tempéré (voir annexe B, § 1) et considérant que la largeur augmente dans les basses fréquences (zone des fondamentaux musicaux), on retiendra que les bandes critiques ont une largeur de l'ordre d'une tierce mineure (18 %), soit un quart d'octave.

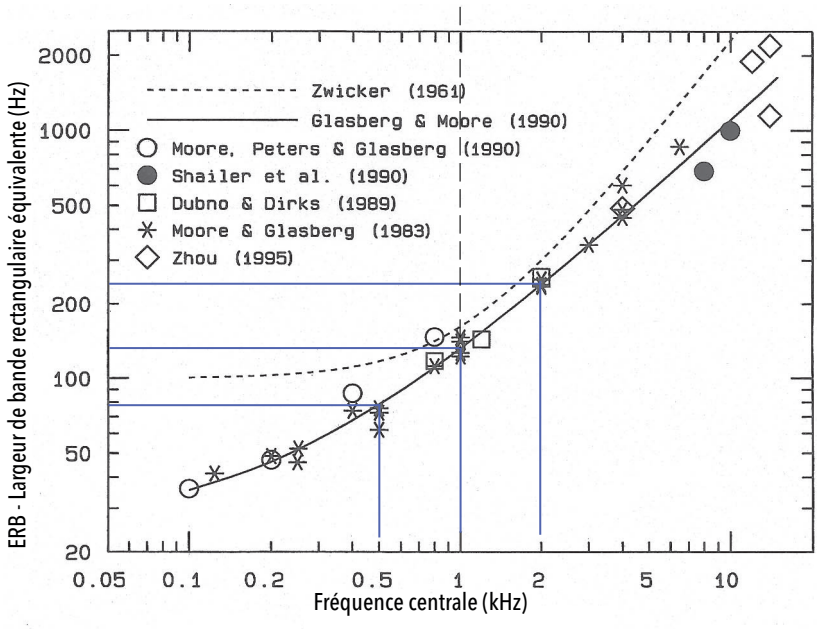


Figure 3.19 Estimation de la largeur des bandes critiques en fonction de la fréquence, d'après les mesures de différents auteurs.

D'après Moore, B., 2013, figure 3.5.

2.2.4. Les bandes critiques et l'analyse fréquentielle

Le concept de bandes critiques est le fondement des théories tonotopiques de la perception de la hauteur. Le principe en est le suivant : lorsque deux sons purs sont émis simultanément, seuls les sons séparés par un intervalle supérieur à celui d'une bande critique (une tierce mineure) sont différentiables à l'oreille. On dit qu'ils sont « résolus » sur la membrane basilaire. Les groupes de cellules de chaque bande répondent aux fréquences de chacun des sons. À l'opposé, deux sons proches dont l'intervalle est inférieur à la tierce mineure ne sont pas résolus. En tombant dans la même bande, ils se combinent. Les groupes de cellules de la bande répondent alors à la modulation d'amplitude résultant de leur combinaison.

Il faut se représenter les filtres auditifs comme une fonction de pondération qui caractérise la sélectivité en fréquence autour d'une fréquence centrale particulière. Il est commode d'utiliser leurs équivalents rectangulaires (ERB *Equivalent Rectangular Band*). À intensité modérée, le filtre auditif est grossièrement symétrique sur une échelle de fréquence logarithmique, de part et d'autre de la fréquence centrale. À plus haut niveau, la pente située du côté des basses fréquences est moins raide que celle qui s'étend du côté des hautes fréquences.

2.2.5. Les phénomènes de masquage

Le masquage est une technique d'exploration des caractéristiques du système auditif qui a donné lieu à un grand nombre d'expérimentations avec des sons sinusoïdaux et des bandes de bruit.

Lors du masquage simultané d'un son par un autre, on peut faire varier le niveau sonore ou la largeur de bande en fréquence du son masquant, pour tester les capacités de discrimination en intensité et en fréquence. C'est ainsi qu'a été défini le *pattern d'excitation*, caractérisé par la sortie de chaque filtre auditif en fonction de sa fréquence centrale. Ces expériences constituent une autre voie d'approche du concept de bande critique. Le test du Son **3.4** (voir encadré page 102), que nous encourageons le lecteur à écouter avec un casque, permet de constater un résultat intéressant : les sons graves masquent facilement les sons aigus.

Deux exemples de masquage

Son 3.4a* (30")

Masquage par un son plus aigu

Son 3.4b* (31")

Masquage par un son plus grave (voir le détail du test page 135)

Dans le domaine temporel, les expériences produisant un masquage, avant ou après un signal, ont pour but de préciser les durées de récupération du système sensoriel après une excitation. Ces expériences mettent en jeu des réactions complexes dont l'interprétation est difficile.

Au-delà des résultats issus de l'expérimentation psychoacoustique, un corpus de connaissances sur le masquage réciproque des sons d'instruments de musique dort dans les traités d'orchestration, et pourrait faire l'objet d'études enrichissantes dans le cadre d'une approche cognitive de la perception sonore.

2.3. La sensibilité aux variations d'amplitude, de fréquence et de durée

Il s'agit d'étudier les réponses d'un sujet aux variations de grandeur d'un stimulus : variations d'amplitude, de fréquence ou de durée, et d'établir des lois de correspondance entre la mesure physique et les réponses d'un groupe de sujets. La valeur moyenne résultant de plusieurs expériences est considérée comme représentative du comportement psychoacoustique humain.

Il n'existe pas en français[10] de termes réservés aux grandeurs perceptives. Nous parlerons donc de sensation d'intensité, de sensation de hauteur et de sensation de durée. Chacun de ces attributs est un paramètre qui a fait l'objet d'expérimentations, afin d'étudier les variations de la sensation selon une seule dimension : par exemple du plus fort au plus faible pour la sensation d'intensité.

Le signal de choix de la psychoacoustique est le son sinusoïdal ou son pur, parfaitement défini par sa fréquence, son intensité et sa durée. Les sons qualifiés de complexes sont considérés comme étant composés de sinusoïdes de fréquences différentes.

Pour « mesurer » les réponses d'un sujet, on procède par comparaison. Les stratégies proposées sont :

- la tâche de *discrimination*. Étant donné deux sons, dire s'il y a ou non une différence entre eux. La tâche de discrimination permet d'estimer les seuils différentiels, c'est-à-dire nos capacités à percevoir les plus infimes changements dans les sons,

10. En anglais, *loudness* et *pitch* sont les termes désignant les sensations dues aux variations d'intensité, de fréquence.

Les sons graves masquent plus aisément les sons aigus

Le masquage est d'expérience courante. À l'écoute d'une multiplicité de signaux sonores simultanés, certains émergent et d'autres sont indiscernables malgré tous nos efforts pour les extraire du bruit ambiant. Lorsqu'un son est masqué, plusieurs paramètres entrent en ligne de compte : la fréquence, la composition spectrale, l'intensité et les caractéristiques temporelles d'occurrence et d'évolution. Le plus souvent, les signaux sont partiellement masqués : ils émergent à des degrés divers. Les signaux d'avertissement sont conçus de telle sorte que leur émergence est assurée quelles que soient les conditions environnantes.

Les expériences psychophysiques de masquage sont réalisées avec des signaux dont on peut contrôler totalement le spectre de fréquence et l'intensité et qui sont dépourvus de sens, afin d'éliminer les aspects cognitifs variables selon les sujets. Ce sont des sinusoïdes et des bandes de bruit. La figure 3.20, ci-contre, montre l'interprétation, sous forme de *pattern* d'excitation, d'une expérience classique du masquage d'un son pur par un bruit. On voit, sur la partie supérieure, les spectres des deux stimuli. Au-dessous sont schématisées les interactions entre les *patterns* d'excitation des deux sons qui produisent un masquage partiel ou total du son pur. La zone ombrée représente la partie du *pattern* d'excitation du son pur qui se trouve éliminée par la présence du bruit. Le bruit ne peut éliminer l'excitation due au son pur que dans la zone d'excitation où son propre *pattern* présente un niveau supérieur à celui du son pur.

Le Sons 3.4a et b font entendre une expérience de masquage similaire faite avec deux sons purs.

On retiendra principalement de ces expériences que les sons aigus sont plus aisément marqués que les sons graves**.** On le comprend on considérant le profil du *pattern* d'excitation qui est dissymétrique : beaucoup plus étendu du côté des hautes fréquences.

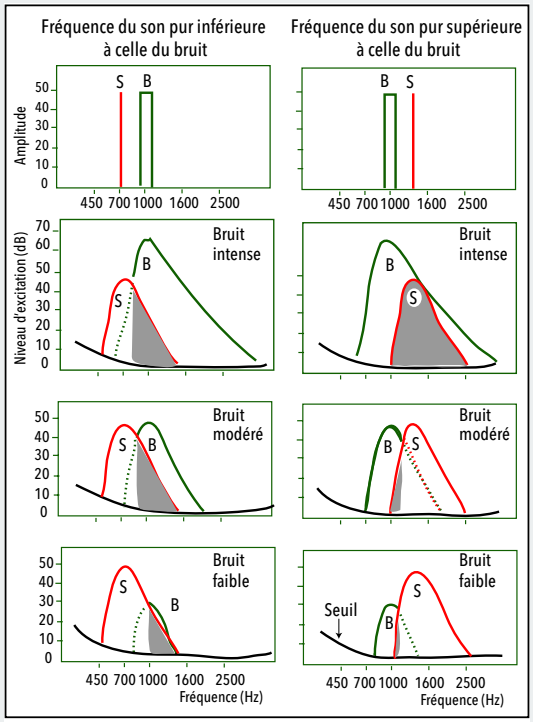


Figure 3.20 *Patterns* d'excitation d'un bruit de bande étroite (B) centré à 1000 Hz et d'un son pur (S) – 700 Hz (à gauche), 1500 Hz (à droite) – pour différents niveaux du bruit. La zone masquée est ombrée.

D'après Botte, M.-C., 1989, figure 1.16 (selon Scharf, 1964, *Acustica* 14).

- la tâche *d'évaluation*. Étant donné un son, ajuster le paramètre d'un deuxième son de façon qu'il paraisse identique au premier, ou qu'il en diffère dans un rapport donné : double, triple, etc. La tâche d'évaluation a permis de construire des échelles psychophysiques pour lesquelles ont été définies des unités spécifiques.

Une des difficultés expérimentales de la psychoacoustique consiste à faire varier un paramètre de façon indépendante des deux autres. Dans la pratique, il est parfois

difficile de savoir lequel on teste effectivement. Nous étudierons tout d'abord les seuils différentiels, ensuite la sensation d'intensité et la sensation de hauteur. Dans tous les cas, ce que les tests permettent d'estimer, ce sont des rapports – c'est-à-dire des intervalles – et non des grandeurs absolues.

2.4. Réponse à l'accroissement d'une stimulation : la loi de Fechner-Weber

Les chercheurs en psychophysique se sont attachés à caractériser la réponse d'un système sensoriel donné à la variation de grandeur du stimulus. La première constatation fut que, soumis à une variation continue de l'excitation, un organe sensoriel fournit une réponse discrète constituée d'une succession de paliers (figure 3.21).

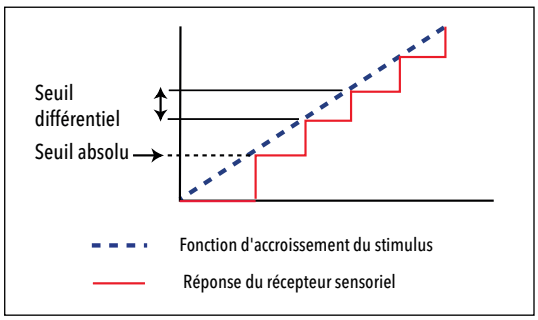


Figure 3.21 Schéma figurant la réponse sensorielle à une excitation dont l'accroissement est continu.

On appelle *seuil différentiel* la plus petite valeur d'accroissement du stimulus qui est juste perceptible.

Comment se comporte le seuil différentiel lorsque la valeur de référence du stimulus change ? Au XVIIIe siècle, l'astronome Bouguer a le premier constaté qu'une différence d'éclairement juste perceptible, correspondant donc au seuil différentiel, avait une valeur relative constante pour différents éclairements[11].

Près d'un siècle plus tard, la loi de Bouguer fut retrouvée par le physiologiste Weber dans un domaine sensoriel tout différent, celui des sensations de pression et de l'appréciation des poids. Cette loi constitua, pour Fechner (1860), la base fondamentale de la psychophysique. Sous sa première formulation, il l'exprime ainsi : « La sensibilité aux différences relatives d'excitation reste constante, quelle que soit la grandeur absolue des excitations. » La relation de proportionnalité qui en rend compte est une fonction logarithmique qui s'écrit :

$$S = A \log E$$

où S est l'accroissement de la sensation, E celui de l'excitation et A une constante de proportionnalité. Dans la pratique, les logarithmes décimaux conviennent pour l'estimation.

On retiendra l'expression : *la sensation croît comme le logarithme de l'excitation*.

La nouveauté du postulat fechnérien, c'est l'idée *d'unités de sensation*. Chaque fois qu'une différence est perçue, un échelon de plus est franchi, d'une unité ayant valeur constante.

La relation logarithmique entre le stimulus et la réponse rend bien compte de la seule donnée importante, qui est le rapport d'estimation entre deux grandeurs,

11. Les données de ce paragraphe sont issues de Piéron, 1945, p. 315, et de l'article de S. Nicolas, 2002.

quelles que soient les quantités mises en jeu. Quand on double la fréquence d'un son pur *f* en passant à 2*f*, 4*f*, l'intervalle perçu est toujours une octave, que l'on passe de 100 à 200 Hz (écart de 100 Hz) ou de 1000 à 2000 Hz (écart de 1000 Hz). De la même façon, quand on double l'intensité sonore en passant de (I) à 2(I), l'accroissement perçu qui est de 3 dB paraît constant à l'écoute.

2.5. Les seuils différentiels d'intensité et de fréquence

Les tableaux 1 et 2 de l'encadré ci-dessous présentent quelques valeurs de seuils différentiels en intensité et en fréquence, pour différents niveaux d'intensité et deux fréquences d'exploration provenant d'une expérience systématique menée aux Bell Telephone Laboratories, et publiée par Stevens et Davis en 1938. La totalité des résultats est représentée sous forme graphique sur les figures 3.22 et 3.23. Les deux graphiques appellent plusieurs remarques.

2.5.1. Les valeurs des seuils différentiels d'intensité et de fréquence

Pour les deux figures les valeurs des seuils différentiels $\Delta x/x$ sont représentées sur une échelle logarithmique. Sur l'ordonnée de droite, les intervalles sont exprimés en cents pour le seuil différentiel de fréquence, et en dB pour le seuil différentiel d'intensité.

Détermination d'un seuil différentiel

Plusieurs méthodes peuvent être employées. Voici un exemple de détermination de seuil différentiel d'intensité. L'expérience se déroule dans une pièce anéchoïque, en audition binaurale. Commençons par la fréquence 1000 Hz. Ayant déterminé l'intensité minimale d'audition de ce son, on émet 1000 Hz à +5 dB au-dessus du seuil : c'est le son A, point de départ du test. Le sujet doit ensuite comparer A et (A + 0,5 dB). S'il n'entend aucune différence, on augmente l'écart en présentant A et (A + 1 dB), ensuite A et (A + 1,5 dB). Lorsque A et (A + Δi) sont perçus comme différents, la valeur Δi est la valeur du seuil différentiel d'intensité à 1000 Hz pour le niveau 5 dB. Dans le tableau 1, le seuil de 1000 Hz à 5 dB est de 3,03 dB. Il faut ensuite recommencer pour le niveau de 10 dB, celui de 15 dB, et ainsi de suite jusqu'au maximum. On procède de la même façon pour les autres fréquences.

Tableau 1. Valeurs de quelques seuils différentiels en intensité di = Δi/i (dB) (Stevens, 1938)

	5 dB	10 dB	20 dB	40 dB	60 dB	80 dB	100 dB
70 Hz	5,74 dB	4,22 dB	2,38 dB	1,04 dB	0,61 dB	---	---
1000 Hz	3,03 dB	2,35 dB	1,46 dB	0,72 dB	0,41 dB	0,29 dB	0,25 dB

Tableau 2. Valeurs de quelques seuils différentiels en fréquence d*f* = Δ*f*/*f* (Stevens, 1938)

	5 dB	10 dB	15 dB	20 dB	40 dB	60 dB	80 dB
62 Hz	0,0975	0,0678	0,0546	0,0491	0,0426	0,0346	---
1000 Hz	0,0094	0,0061	0,0044	0,0039	0,0036	0,0034	0,0030

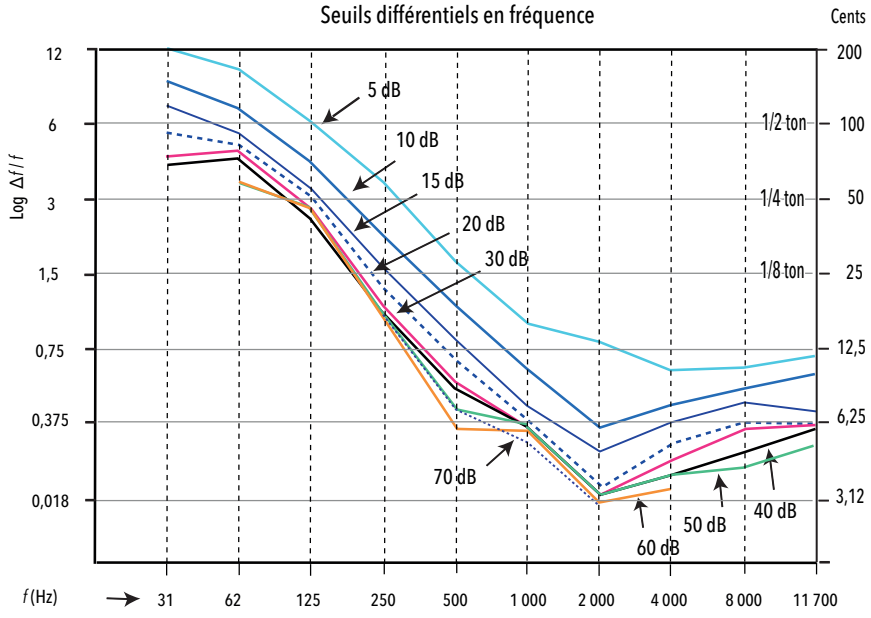


Figure 3.22 Valeur des seuils différentiels pour différentes fréquences. Les intervalles sont exprimés en pourcentage de la fréquence (Log $\Delta f/f$) et en cents.

D'après les données de Stevens, S., & Davis, H., 1938, pages 140.

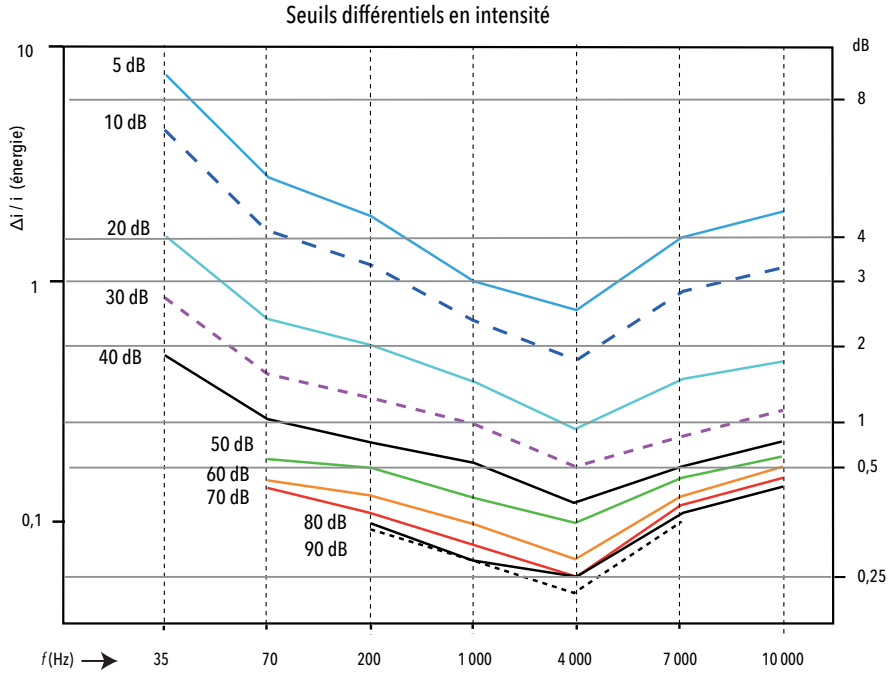


Figure 3.23 Valeur des seuils différentiels pour différentes intensités sonores. Les intervalles sont exprimés en pourcentage de l'énergie ($\Delta i/i$) et en décibels (dB).

D'après les données de Stevens, S., & Davis, H., 1938, pages 89.

Si le seuil différentiel était constant, tous les points seraient sur une droite horizontale. On voit qu'il n'en est rien. Pour la fréquence comme pour l'intensité, le seuil différentiel diminue rapidement lorsque le niveau d'écoute augmente, puis se stabilise plus ou moins lorsque l'intensité moyenne est de 80 dB. Par ailleurs, les valeurs du seuil sont plus grandes du côté des basses fréquences (< 1000 Hz). Elles diminuent lorsqu'on s'approche de la zone de plus grande sensibilité de l'oreille, entre 2000 et 4000 Hz, zone à l'intérieur de laquelle on rencontre les plus petites valeurs.

Il faut retenir que les seuils différentiels ne sont pas constants : ils varient avec l'intensité et avec la fréquence, ils diminuent lorsque l'intensité du son augmente, et les plus petites valeurs des seuils se rencontrent dans la zone de fréquence où l'oreille est la plus sensible, entre 1000 et 4000 Hz.

Les deux tests (Son **3.5** et Son **3.6**), à écouter au casque, permettent d'estimer ses propres seuils différentiels en fréquence (à 1000 Hz) et en intensité (à 800 Hz).

Son 3.5* (33")

Seuil différentiel d'intensité (10 sons tests). Voir légende du son page 136.

Son 3.6* (2'11)

Seuil différentiel de fréquence (10 groupes de sons). Voir la légende du son page 136.

Les seuils différentiels varient selon les sujets et les conditions expérimentales. Les mesures présentées figures 3.22 et 3.23 ont été faites avec des sujets n'ayant aucune compétence particulière pour la musique[12]. Or, les plus faibles valeurs sont remarquablement fines puisqu'on atteint 1/4 de décibel à 4000 Hz pour 80 dB, et 3 cents à 2000 Hz pour 60 dB (le cent est une unité d'intervalle qui vaut 1/100 de demi-ton tempéré). Il peut sembler paradoxal que le même récepteur, en l'occurrence l'oreille, nous permette de capter une large gamme dynamique – puisque entre le seuil d'audibilité et le seuil de douleur nous acceptons une variation de 10^{1} à 10^{12} soit 120 dB – et que par ailleurs nous soyons capables d'apprécier d'aussi faibles variations que le 1/4 de dB. De la même façon, il est étonnant de penser que nous pouvons percevoir un intervalle aussi faible que le 3/100 de 1/2 ton, alors que notre musique admet comme plus petit intervalle le demi-ton, soit un intervalle 30 fois plus « grand ». Mais il y a lieu de faire une distinction entre les catégories d'intervalles, en fréquence ou en intensité, constitutives de la structure d'une musique, et la perception des écarts par rapport à ces catégories, qui sont une des sources du plaisir de l'écoute. Il ne s'agit plus de physiologie de la perception mais de stratégies cognitives. Nous y revenons au chapitre 8 à propos des systèmes d'accordage des instruments de musique. En d'autre termes, les seuils différentiels attestent une fois de plus que nos sens peuvent capter de façon très fine les faibles variations temporelles de l'intensité et de la fréquence des sons, celles qui font la « vie des sons réels », sans pour autant altérer l'identité catégorielle de ces phénomènes.

2.5.2. La carte de discrimination auditive de l'oreille

La figure 3.24 reprend celle que Stevens et Davis proposent dans leur ouvrage (1938). Elle représente un calcul estimatif de *quanta* perceptifs, qui est le produit des seuils différentiels d*f*.di (voir encadré page 104), estimés sur l'aire audible quadrillée selon un maillage de demi-octaves en fréquence et de 10 dB en intensité. Nous proposons figure 3.25 une interprétation de ces résultats sous forme d'une carte de niveaux de discrimination auditive (voir aussi Xenakis, 1960, page 66).

Il est clair que c'est dans la zone centrale G, zone sensible de l'oreille et à fort niveau d'écoute, que le nombre total de sons discriminables est le plus élevé. De part et d'autre, vers les hautes fréquences et vers les basses fréquences, la capacité de discrimination diminue, et s'affaiblit aussi avec l'intensité. En ce qui concerne la pratique musicale, on peut faire deux remarques. D'une part, la grande sensibilité aux fines variations mélodiques dans le registre aigu des instruments et de la voix (500 à 2000 Hz), d'autre part, l'importante contribution des harmoniques dans

12. Douze sujets pour le test du seuil différentiel en intensité et cinq hommes (20 à 30 ans) pour le test du seuil différentiel en fréquence.

l'appréciation fine des variations d'intensité et de sonorité. La séquence musicale des Sons **3.7** et **3.8** dans laquelle l'exposition d'un thème est d'abord présentée par les instruments les plus graves offre une belle illustration de la première remarque.

Son 3.7 (10")

Octuor : thème joué au violoncelle (2 fois)

Son 3.8 (21")

Octuor : entrée successive des instruments, du grave à l'aigu

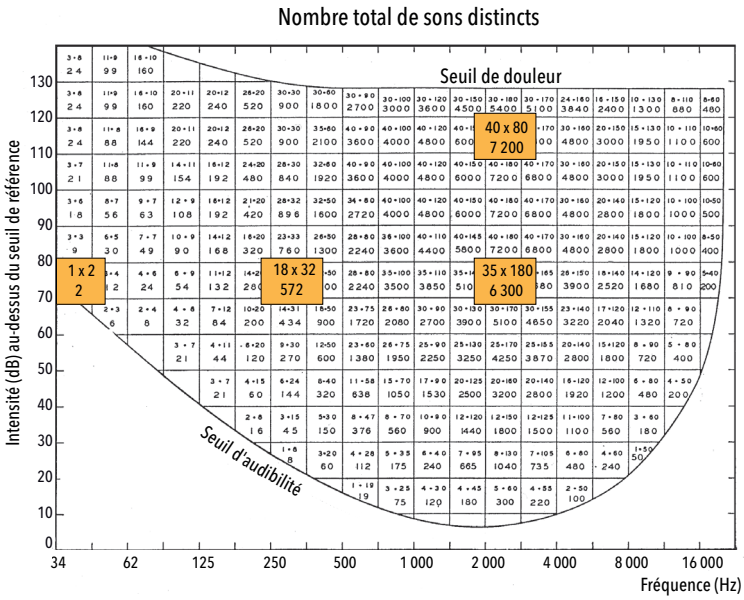


Figure 3.24 Calcul des « quantas » acoustiques sur la base des seuils différentiels.

D'après Stevens, S., & Davis, H., 1938, p. 153.

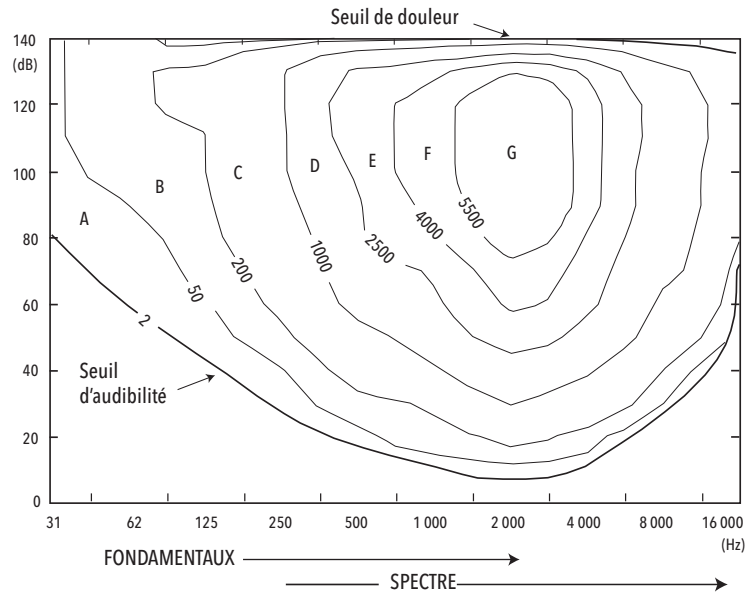


Figure 3.25 Représentation schématique de la figure 3.24 sous forme de courbes de niveau. La zone de discrimination la plus fine correspond à G (2000-4 000 Hz, intensités élevées).

Castellengo, M., 1994, *La perception auditive des sons musicaux*, figure 2.

2.6. La sensation d'intensité des sons purs

La sensation d'intensité, liée à l'amplitude des vibrations sonore, se présente comme une aptitude banale que chacun possède naturellement et qui, contrairement à la sensation de hauteur, ne nécessite pas d'apprentissage particulier. Pourtant elle pose de difficiles problèmes de mesure.

2.6.1. La variation de l'intensité avec la fréquence : courbes d'égale sensation d'intensité sonore

La première exploration consiste à évaluer l'égalité d'intensité sonore ou *isophonie.* Il s'agit d'établir des courbes similaires à celle du seuil d'audibilité (voir page 95), mais pour différents niveaux de l'intensité du son.

La fréquence de 1000 Hz étant prise comme référence, on pose que l'intensité perçue d'un son de 1000 Hz, exprimée en *phones*, a la même valeur que sa mesure en dB. L'intensité perçue d'un son de 1000 Hz à 40 dB vaut donc 40 phones. Pour tracer la courbe d'égale sensation sonore à 40 phones pour toutes les fréquences, on recherche, pour chacune d'elles, la valeur de l'intensité en dB qui produit une sensation d'intensité équivalente. Depuis Fletcher et Munson (1933), l'expérience a été répétée un grand nombre de fois et les courbes obtenues ont fait l'objet d'une norme internationale[13].

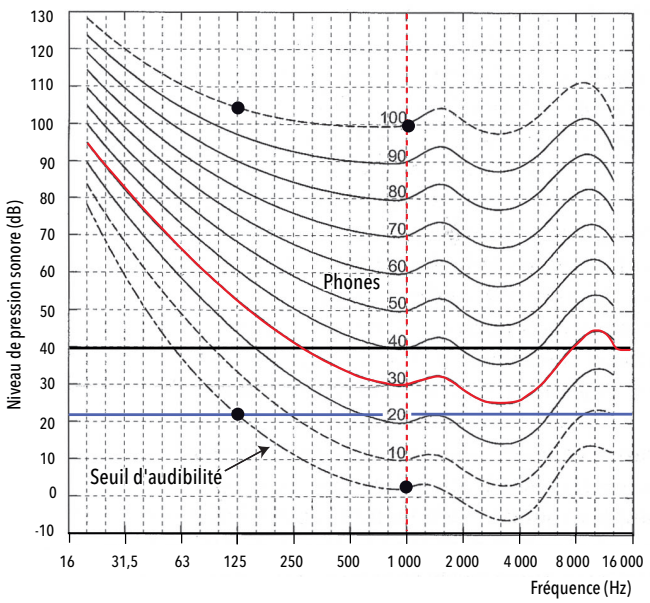


Figure 3.26 Courbes d'égale sensation d'intensité sonore (sons purs). Norme ISO 2003. Son frontal, mesure binaurale, en champ libre. Les résultats expérimentaux correspondent aux courbes en trait continu.

Moore, B., 2013, figure 4.1.

On voit sur la figure 3.26 que les courbes d'isophonie suivent d'assez près celle du seuil d'audibilité, mais tendent à s'aplatir avec l'augmentation d'intensité. Prenons un exemple. À 125 Hz (do_2), le seuil d'audibilité est à 22 dB, soit 20 dB au-dessus de celui du son de 1000 Hz qui est à 2 dB (axe bleu). À 100 phones, le son de 125 Hz est à 104 dB et celui de 1000 Hz à 100 dB : la différence n'est plus que de 4 dB.

13. Norme ISO dont la dernière publication date de 2005.

Corrélativement, la dynamique du son de 125 Hz est de 82 dB, contre 98 dB pour le son de 1000 Hz. Les différences d'intensité perçues entre sons graves et sons aigus sont beaucoup plus marquées aux faibles niveaux sonores. On peut en faire l'expérience en écoutant le Son **3.9** dans lequel trois sons purs, dont la fréquence est de 100 Hz, 1000 Hz et 10000 Hz, sont produits à la même intensité physique. Il faut tout d'abord régler le niveau d'écoute de sorte que le premier son, 100 Hz, soit tout juste audible, ensuite écouter 1000 Hz et 10000 Hz. Le son de 1000 Hz paraît nettement plus intense que les deux autres. On pourra répéter l'expérience à un niveau plus fort et constater que la perception des différences d'intensité entre les sons s'atténue. La variation de la dynamique perçue selon les fréquences a pour conséquence un changement d'équilibre entre sons graves et sons aigus lorsqu'on change le niveau d'écoute d'une musique. Certains appareils possèdent une position *loudness* qui introduit un filtre correcteur des basses fréquences. La courbe de 30 phones est prise comme référence pour la normalisation de la mesure de l'intensité en dBA.

Son 3.9 (11")

Trois sons purs (il faut de bonnes enceintes !)

On voit quelquefois, sur les courbes d'isophonie, les indications de nuances *pp, mf, f, fff,* ce qui porte à croire que l'on pourrait utiliser un appareil gradué en phones pour mesurer les nuances musicales. C'est une utopie. Il ne faut pas perdre de vue que ces courbes sont établies pour des sons purs, alors que les sons instrumentaux sont riches en harmoniques, et que leur spectre varie de façon considérable avec l'intensité (voir chapitre 5, § 3, page 210).

2.6.2. L'appréciation des intervalles d'intensité : l'échelle des sones

S'il est relativement aisé d'ajuster l'égalisation d'intensité sonore entre deux sons de même fréquence, ou de les ordonner par intensité croissante ou décroissante, il n'en est pas de même pour l'établissement d'une échelle quantitative. La question qui se pose est : peut-on ajuster l'intensité d'un son de sorte qu'il soit perçu deux fois ou trois fois plus fort que sa valeur initiale ? À la suite d'expériences réalisées en 1957, Stevens conclut que le son paraît deux fois plus fort lorsque l'intensité sonore est incrémentée de 10 dB (voir Moore, page 131). Il définit alors une nouvelle unité, le *sone*, toujours basée sur la fréquence 1000 Hz. Par convention, 1 *sone* correspond à l'intensité perçue d'un son de 1000 Hz et de 40 phones présenté dans les conditions suivantes : incidence frontale, écoute binaurale, en champ libre, dans une pièce anéchoïque. De même, 2 sones correspondent à l'intensité perçue d'un son de 1000 Hz à 50 phones, 4 sones à 60 phones, 8 sones à 70 phones[14]. Les expériences de Stevens ont été reprises par plusieurs auteurs, mais les résultats montrent une assez grande variabilité selon les sujets et selon les protocoles. De fait, il est extrêmement difficile d'apprécier quantitativement l'intensité de sons purs de fréquence différente et, comme nous le verrons dans le chapitre 5, d'autres indices que la mesure de l'énergie nous renseignent ordinairement sur l'intensité des sons.

2.6.3. La sensation d'intensité et la durée des sons

La durée a une incidence notable sur la perception d'intensité des sons. Il est difficile de mesurer l'intensité perçue de sons brefs, par exemple, des sons de 2 à 3 dixièmes de seconde. Aux faibles niveaux, proches du seuil d'audibilité, l'augmentation de durée améliore la détection, ce qui peut être interprété soit comme une indication que l'oreille intègre l'énergie, soit comme une augmentation de la probabilité de détection.

14. On a ainsi log (S) = 0,03 (Lph – 40) où « S » est la valeur cherchée en sone ; « Lph » l'intensité en phones (*Wikimedia, article Fundamentals of Psychoacoustics*).

2.7. La sensation de hauteur des sons purs

La sensation de hauteur d'un son pur est directement donnée par sa fréquence fondamentale, mais, comme on l'a vu pour l'intensité, la relation n'est pas indépendante des autres paramètres.

2.7.1. Les variations de la sensation de hauteur avec l'intensité

Dès qu'il a été possible de produire des sons sinusoïdaux de fréquence et d'intensité variable, les expérimentateurs ont remarqué que la sensation de hauteur évoquée par un son pur changeait avec l'intensité, mais différemment selon la fréquence. Les sons graves, inférieurs à 500 Hz, semblent baisser avec l'augmentation d'intensité alors que les sons aigus, supérieurs à 3000 Hz, paraissent monter. Stevens a proposé une nouvelle unité, *le mel*, qui tient compte de l'effet de l'intensité sur la hauteur perçue, en fonction de la fréquence. Cependant, l'effet est très variable selon les auditeurs. De plus, pour un même auditeur, la sensation peut changer selon qu'il fait l'expérience avec l'oreille gauche ou avec l'oreille droite. L'incidence de l'intensité sur la sensation de hauteur est négligeable dans la pratique musicale puisque les instruments traditionnels sont riches en harmoniques, mais elle doit être prise en considération lorsqu'il s'agit de musiques de synthèse.

2.7.2. Les variations de la sensation de hauteur avec la durée : seuil temporel

Son 3.10* (19″)

Seuil à 300 Hz

Son 3.11* (22″)

Seuil à 1000 Hz

Son 3.12* (18″)

Seuil à 3000 Hz

Il faut une durée minimale pour percevoir la fréquence d'un son pur. L'expérience proposée (Sons **3.10** à **3.12)** donne à entendre, pour trois fréquences différentes, des portions de sinusoïdes contenant un nombre croissant de périodes : 1, 2, 4, 8, etc.

Les premiers sons de chaque exemple donnent la sensation d'un clic ou d'un bruit coloré très bref dont la hauteur se précise plus ou moins rapidement selon les auditeurs.

L'analyse sonagraphique de la figure 3.27 fournit une bonne représentation de ce qui semble se passer dans le système auditif, lequel n'échappe pas au principe d'incertitude déjà évoqué : $\Delta f \times \Delta t$ = constante. Comme on le ressent généralement à l'écoute, la hauteur perçue au cours de cette succession de sons de durée croissante varie, mais selon les sujets, elle monte ou descend graduellement. Pour les sons très courts, nous serions donc moins sensibles à la fréquence centrale qu'à celle de l'un des bords de la bande de bruit, dont la largeur diminue au fur et à mesure que la durée du son croît (voir pointillés sur la figure).

Pendant longtemps, le nombre de sons qu'il était possible d'articuler dans une seconde a été limité par les capacités gestuelles humaines. Si l'on envisage 12 à 15 notes par seconde, ce qui semble une limite[15], et qu'on suppose les notes égales, la durée unitaire oscille de 60 à 80 ms, mais celle de la partie réellement entretenue du son est en fait plus réduite en raison des transitoires. La discrimination de sons brefs consécutifs est meilleure dans l'aigu de la tessiture puisque, pour une durée donnée, le son contient un plus grand nombre de périodes. Ces limites ont été rapidement dépassées dès qu'il est devenu possible de manipuler des sons enregistrés sur bande, mais alors les auteurs se sont vite aperçus qu'une autre donnée était à prendre en compte : celle de l'ordre des sons dans la séquence[16].

15. Nous ne prenons pas en compte ici le jeu *glissando* sur un clavier, par balayage des touches avec un doigt.
16. Voir à ce sujet l'article de G. Ligeti, « Musique et technique, Expériences personnelles et considérations subjectives », 2001.

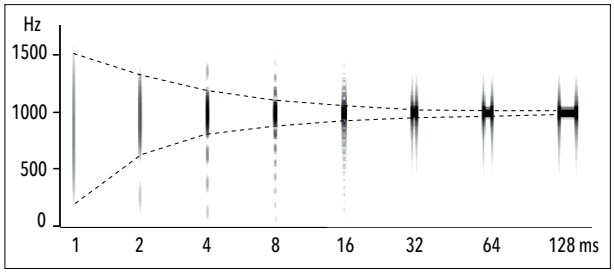


Figure 3.27 Analyse du Son 3.11. Présentation d'un son de 1000 Hz dont la durée, initialement d'une période, est doublée à chaque présentation. Analyse FFT ; fenêtre Hanning ; 2 048 pts à 44,1 kHz, donc $\Delta f = 107$ Hz et $\Delta t = 46$ ms.

2.7.3. Les échelles de sensation de hauteur : intervalles de hauteur

La dimension de hauteur se prête aisément à la constitution d'échelles et à la définition d'intervalles. Parmi eux l'octave tient une place privilégiée[17].

Cette constatation a donné lieu à la conception d'une échelle des hauteurs que l'on représente sous forme d'une hélice ascendante dans laquelle le parcours d'un tour correspond à l'intervalle d'octave. Ce concept permet de combiner deux aspects de la sensation de hauteur : la *hauteur brute*[18], qui est en relation directe avec la fréquence du son, depuis les plus faibles valeurs jusqu'aux plus grandes, et le *chroma* qui correspond à la progression des intervalles à l'intérieur d'une octave. Le chroma se reproduit à l'identique aux différents étages de l'hélice. Une telle conception rend compte des sensations développées à l'écoute de sons purs. Les sons purs de fréquences différentes changent de qualité sonore avec la tessiture : on reconnaît les graves par la sensation de bourdonnement doux, les médiums à leur clarté et les aigus à leur caractère perçant. En ce qui concerne l'évaluation des intervalles, les musiciens sont souvent désorientés par l'étrangeté du son sinusoïdal dans lequel ils ne retrouvent pas leurs repères habituels.

Le rapport de fréquence entre deux sons à intervalle d'octave est de 2. Ainsi la_2, la_3, la_4 ont respectivement pour fréquence 220 Hz, 440 Hz et 880 Hz. Toutefois, la correspondance entre mesure physique et perception n'est pas valable sur toute la tessiture. Le Son **3.13**, qui donne à entendre une suite de sons purs dont la fréquence double à chaque présentation, permet de faire soi-même l'expérience, dans l'intervalle 62,5 Hz (do_1) à 16 000 Hz (do_9).

Son 3.13* (24")

Justesse des octaves : huit couples de sons (voir la légende du son page 136)

Bien que les réactions diffèrent selon les sujets, on peut dégager les tendances suivantes :

- dans la zone médium, entre 125 et 1000 Hz, la mesure physique correspond généralement bien à la perception. Les sons dont les fréquences sont doubles l'une de l'autre donnent la sensation d'octaves justes ;
- aux fréquences inférieures, entre 60 et 125 Hz, les réponses des auditeurs divergent. Pour certains, l'octave paraît trop grande, pour d'autres, elle est trop courte ou juste ;

17. Voir le chapitre 8 pour une étude plus détaillée de la relation d'octave et de son rôle dans la constitution des échelles musicales.
18. Dans le cadre plus général des sons complexes réels, c'est ce que nous désignons par le terme de hauteur spectrale.

- à partir de 1000 Hz et en montant dans l'aigu, l'intervalle d'octave paraît généralement de plus en plus petit ;
- au-delà de 4000 Hz, les réponses deviennent erratiques : les auditeurs ont du mal à porter un jugement.

Cette expérience appelle plusieurs remarques. Tout d'abord, la zone de bonne concordance entre le rapport de fréquence et l'estimation d'octave correspond à celle de la plupart des fondamentaux musicaux, c'est-à-dire les sons à l'intérieur des portées de clé de sol et clé de fa. Par ailleurs, les effets observés dans la tessiture grave, nets avec des sons sinusoïdaux, disparaissent lorsqu'on emploie des sons riches en harmoniques. Enfin, le fait que les intervalles physiquement corrects paraissent trop petits dans l'aigu semble une particularité physiologique indépendante du contenu du son, que nous retrouverons à propos de l'accordage des instruments de musique dans l'aigu.

2.7.4. La limite aiguë de la sensation d'intervalle

Chaque son dure 4".

Il faut impérativement commencer par le son 3.14a.

Son 3.14a

Son 3.14b

Son 3.14c

Son 3.14d

Son 3.14e

Voir les commentaires page 137.

Nous avons déjà fait état, à propos du codage temporel (voir § 1.5.3), d'une limite supérieure du codage de la fréquence, limite estimée aux alentours de 4 à 5 kHz. Il s'agit, comme toujours, d'une valeur moyenne, variable selon l'expertise des sujets. Le mieux est de faire soi-même l'expérienc en se prêtant à l'écoute d'un test (Sons **3.14 a** à **e**). La première séquence entendue évolue dans une tessiture extrêmement aiguë (voir les commentaires du son 3.14 page 137). Quelques (rares) personnes entendent précisément une mélodie à cette hauteur, et peuvent la noter. La plupart des auditeurs ne peuvent noter que la 2e et surtout la 3e mélodie qui se situent à des tessitures de moins en moins aiguës.

La fréquence 4000 Hz, do_7, limite pratique des instruments mélodiques, se situe dans la dernière octave du piano.

2.7.5. Petit bilan de la perception des sons purs

Pour l'intensité comme pour la fréquence, nous possédons une très grande sensibilité différentielle aux plus infimes changements se produisant dans le signal. Nous pouvons user de cette capacité pour être attentif à ce qui se passe autour de nous. Mais en ce qui concerne l'aspect dimensionnel, c'est-à-dire la définition d'échelles et l'appréciation des intervalles, force est de constater que nous sommes peu armés pour repérer et mémoriser les variations quantitatives d'intensité, car elle change en permanence, selon la distance à la source et les propriétés acoustiques du lieu. Au contraire, la fréquence n'est que très rarement affectée[19] : c'est un paramètre fiable et reproductible, dont les valeurs sont mémorisables. Elle permet l'établissement d'une échelle et l'appréciation d'intervalles.

Les sons purs sont de bons stimuli pour tester les capacités de l'oreille en tant que récepteur (au niveau physiologique). Cependant, dès qu'il s'agit de tests avec des sujets, il faut prendre garde à l'ennui et à la fatigue qui influent sur les résultats. De surcroît, les musiciens, dotés de très bonnes « oreilles » dans leur pratique sont souvent désarmés à l'écoute de ces signaux étranges que sont les sons purs. Pour la sensation d'intensité comme pour celle de hauteur, on peut établir une relation logarithmique entre l'unité de mesure physique et la réponse sensorielle. Nous verrons au chapitre suivant que les *lois* de comportement, établies pour les sons purs, sont difficilement transposables aux sons réels de notre environnement

19. Citons par exemple l'effet Doppler dû au déplacement de la source.

produits par des sources connues. Les connaissances acquises antérieurement sur le comportement de ces sources sont autant de clés pour l'interprétation des variations de hauteur et d'intensité. Déjà en 1874, Helmholtz remarquait :

> *Dans l'usage que nous faisons des organes de nos sens, l'exercice et l'expérience jouent un rôle beaucoup plus important qu'on n'est ordinairement porté à le croire, et comme ainsi que nous venons de le faire remarquer, nos sensations n'ont d'importance pour nous qu'autant qu'elles nous mettent à même de connaître exactement le monde extérieur, notre attention ne s'applique ordinairement à l'observation de ces sensations, que juste autant qu'il est nécessaire pour atteindre à ce but.*
>
> Helmholtz, H., 1874, p. 82.

2.8. La sensation d'intensité des sons complexes

L'évaluation de la sensation d'intensité est un enjeu important dans les domaines de la mesure du bruit. Les méthodes de calcul de la sonie, dérivées de celle de Zwicker (1999), ont comme point commun le filtrage du spectre selon les 24 bandes critiques (Barks, ou ERBs). Les largeurs de bande et la forme des filtres ont été normalisés de façon à aboutir à une évaluation quantitative pratique pour la réglementation, et sont maintenant intégrées aux logiciels d'usage courant dans l'industrie. Cependant, lorsqu'il s'agit de sons fluctuants en intensité, et plus particulièrement d'impulsions, par exemple les bruits d'explosion d'un moteur de mobylette, les valeurs calculées sont encore loin de la réalité perceptive.

L'application la plus importante du calcul de la sonie est celle du codage numérique et de l'élaboration des différents systèmes de compression des signaux audionumériques. Jointe aux données du masquage, elle permet de supprimer du spectre les fréquences non perçues, autorisant ainsi d'importantes réductions des données tout en conservant une bonne qualité perçue. Diverses sortes de compression se sont développées, depuis l'algorithme ATRAC créé par Sony pour le MiniDisc (taux de compression de 5), jusqu'au MP3 qui conserve une très bonne qualité[20] avec un taux de compression de 10.

2.9. La sensation de hauteur des sons complexes

2.9.1. Du son pur au son complexe

Les sons complexes étudiés en psychoacoustique sont des sons stables composés de plusieurs fréquences pures, en relation harmonique ou non[21]. La hauteur d'un son complexe est mesurée par comparaison avec la fréquence du son pur donnant la même sensation de hauteur.

L'étude des sons complexes en psychoacoustique sert en quelque sorte de pierre de touche pour éprouver la validité des deux théories de perception de la hauteur en présence : la théorie tonotopique (page 91) et la théorie temporelle (page 93). Dans le cas du son pur, la position du maximum d'excitation dans la cochlée et la (ou les) fibre(s) assurant la transmission des influx se correspondent de façon univoque ; seule l'interprétation proposée pour la mesure de la fréquence diffère.

20. À condition de disposer d'un bon codeur et de bien choisir le « débit », classiquement 256 kbits/s. Voir les nombreux documents sur Internet.
21. On remarquera que le critère de distinction qu'opère la psychoacoustique entre sons *simples* (ou purs) et sons *complexes* diffère fondamentalement de celui que nous adoptons pour l'étude de la perception de la hauteur, qui est le mode de production sonore. Nous distinguons les sons entretenus, *périodiques*, qu'ils soient simples ou complexes, et les sons non entretenus, *apériodiques*.

Le spectre d'un son pur n'a qu'une composante, qui est aussi la fondamentale (voir figure 3.28 A). Les sons complexes (figure 3.28 B à E, Son **3.15**), posent de nouveaux problèmes :

Chaque son dure 3".

Son 3.15a

Son 3.15b

Son 3.15c

Son 3.15d

Son 3.15e

Les résultats peuvent être différents selon que vous écoutez au casque ou sur des enceintes.

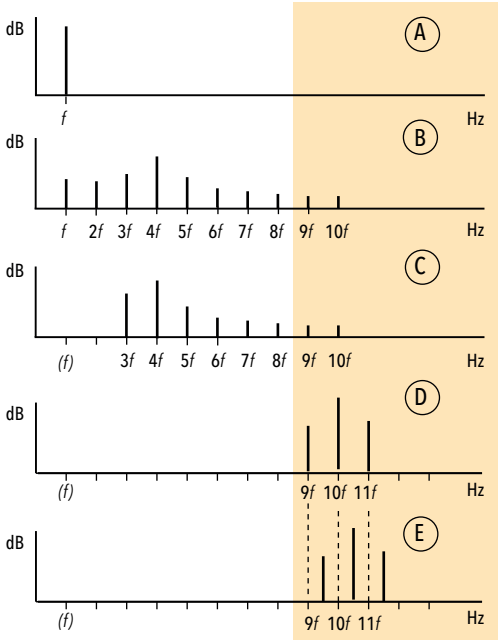


Figure 3.28 Spectres de différents sons complexes. Les composantes des sons A, B, C sont « résolues » (voir § 2.2.4). D et E sont des modulations dont les composantes, non résolues, sont périodiques pour D et apériodiques pour E.

- chacune des fréquences du son complexe pouvant être candidate à la sensation de hauteur, il s'agit d'expliquer la prédominance de la fréquence fondamentale sur celles des autres composantes ;
- le maximum d'énergie du spectre ne coïncide pas toujours avec le fondamental (B) ;
- pour certains sons, seules les composantes de rang élevé existent dans le signal, alors que la sensation de hauteur évoquée correspond à une fréquence plus basse[22], située dans une zone où il n'y a aucune énergie (C et D) ;
- les sons complexes périodiques peuvent donner lieu à la perception de plusieurs hauteurs ;
- les sons apériodiques (E) ont des hauteurs ambiguës.

Deux groupes de théories se proposent d'expliquer la perception de hauteur des sons complexes.

Il s'agit des modèles dits de *reconnaissance de forme*[23], qui nécessitent la présence de composantes *résolues*, c'est-à-dire suffisamment distantes pour affecter des bandes critiques différentes, et des modèles faisant intervenir le codage temporel,

22. C'est plus fréquent que ne le disent les psychoacousticiens. Les causes en sont : le faible rayonnement de l'énergie dans les graves (instruments de petites dimensions), le filtrage ou encore des techniques musicales particulières (chant diphonique, arc musical, guimbarde).
23. Traduction de l'anglais *pattern recognition model*.

qui requièrent l'existence de composantes de rang élevé interférant entre elles à l'intérieur d'une bande critique.

2.9.2. Les théories tonotopiques : modèles de reconnaissance des formes

La perception de la hauteur se déroule en deux étapes. La première est une analyse qui détermine les fréquences de quelques-unes des composantes sinusoïdales isolées du son complexe. Dans la seconde étape, un « reconnaisseur de forme », différent selon les modèles, détermine la hauteur du son complexe sur la base des fréquences des sons isolés. Le codage tonotopique expliquerait le fait que l'on peut séparer à l'oreille les premiers harmoniques d'un son, comme l'ont remarqué en leur temps Mersenne et Descartes :

> *la chorde frappée, & sonnée à vuide fait du moins cinq sons différens en même temps, dont le premier est le son naturel de la chorde... Or il faut choisir un grand silence pour les appercevoir, encore qu'il ne soit plus nécessaire quand on y a l'oreille accoustumée... Quant à moy je n'y ay nulle difficulté, & j'ay rencontré plusieurs Musiciens qui les entendent aussi bien que moy... c'est pourquoy je mets icy les observations que j'ay justifiées très-exactement plus de cent fois...*
>
> Marin Mersenne, 1636, *Harmonie universelle*, Livre quatrième des instruments, Proposition IX.

On remarquera en écoutant l'exemple sonore **3.16** que la capacité à discriminer les harmoniques dépend de la tessiture. Elle est plus aisée dans le grave que dans l'aigu. Voir aussi chapitre 7 l'exemple du jeu de tierce à l'orgue, § 2.4. De façon similaire, les auteurs expliquent la perception de hauteur des sons inharmoniques et les ambiguïtés de perception de la fondamentale observables selon les sujets.

Son 3.16 (18")

Discrimination des harmoniques selon la tessiture. Voir la légende du son page 137.

L'interprétation des phénomènes devient plus complexe lorsque la hauteur perçue est « virtuelle », c'est-à-dire lorsque le son ne possède pas d'énergie à la fréquence fondamentale. Prenons l'exemple du son périodique C (figure 3.28). Le reconnaisseur de forme qui doit trouver cette fréquence fondamentale diffère selon les auteurs. Dans ce qui suit, nous résumons les théories analysées par Brian Moore (2013, chapitre 6, § 4). Walliser (1969) prend en compte le partiel le plus grave 3f dont il décline les sous-multiples : celui dont la fréquence correspond à l'intervalle entre les autres composantes est le fondamental. Terhardt[24] invoque le rôle d'un « partiel dominant » soit par son intensité, soit parce qu'il est situé dans la région fréquentielle comprise entre 500 et 1500 Hz. Il prend en compte également la reconnaissance d'un modèle harmonique. Goldstein ne retient que l'information fréquence. Selon lui le processeur postule que tous les stimuli sont périodiques et trouve la série harmonique qui s'ajuste le mieux à la série de composantes présentes.

Les modèles de reconnaissance de forme sont mis en défaut par la sensation de hauteur que produisent indubitablement des sons ne comportant que des composantes non résolues, c'est-à-dire des sons dont l'intervalle entre les composantes successives est beaucoup plus faible que la largeur d'une bande critique, ce qui, pour un son périodique, correspond aux harmoniques supérieurs à 8 (fig 3.28 D).

24. Terhardt (voir page Web personnelle) a proposé un algorithme de calcul de la hauteur des sons complexes, disponible dans certains logiciels.

2.9.3. Les théories temporelles : le concept de résidu

À l'inverse des modèles précédents, les théories du codage temporel, dont le principal représentant est Schouten, requièrent l'existence de composantes de rang élevé interférant entre elles à l'intérieur d'une bande critique (figure 3.28 D et E). La détection de la période se ferait sur la base de la synchronisation des influx sur les maxima adjacents de la forme d'onde, et non sur la période globale du battement, comme indiqué figure 3.29.

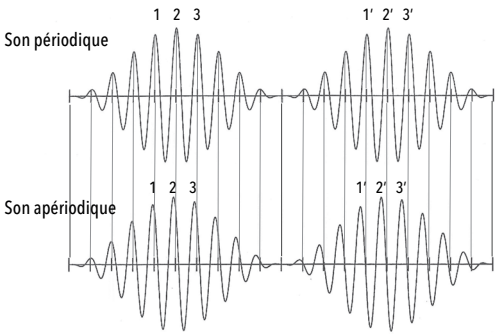


Figure 3.29 Exemple de détection de pics adjacents sur deux signaux modulés en fréquence. En haut : son pur de 2 000 Hz modulé à 200 Hz ; en bas, son pur de 2 030 Hz modulé à 200 Hz.

D'après Moore, B., 2013, figure 6.7.

La hauteur virtuelle : région d'existence et principe de dominance

Plusieurs expériences ont été réalisées pour comprendre la perception des sons ne comportant que des composantes aiguës. Les auteurs ont fait varier le nombre des composantes, leur rang et la zone de fréquence dans laquelle elles se situent. Ils ont mis en évidence certaines particularités, comme le principe de dominance de certains partiels ou la région d'existence du fondamental dont nous verrons tout l'intérêt au chapitre 6.

Ritsma (1962) a montré que le caractère tonal de la hauteur résiduelle perçue ne se produisait que dans l'intervalle de fréquence de 100 à 400 Hz, qu'il a appelé *région d'existence*. Quelques-uns des harmoniques du son doivent aussi être inférieurs à 5 kHz. Il a également formulé (Ritsma, 1967) le *principe de dominance* des composantes de rang 3 à 5 (figure 3.30).

Le poids particulier des harmoniques 3, 4 et 5 dans la détermination de la hauteur perçue a été confirmé par plusieurs auteurs. Mais, si la tessiture des composantes est très aiguë, la dominance se déplace vers les premiers harmoniques. Au-dessus de 1 000 Hz, le fondamental, s'il existe, est dominant.

Ces observations trouvent écho dans l'interprétation de la hauteur de sons inharmoniques comme les cloches (chapitre 6, § 3.4, § 3.5.1, § 3.5.3), les sons multiphoniques (chapitre 8, § 2.4), pour certains jeux de l'orgue (chapitre 7, § 2.4.1) et pour le son virtuel de la « quintina » (chapitre 9, § 3.5).

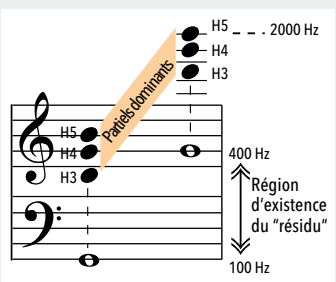


Figure 3.30 Région d'existence du « résidu » et zone d'existence des partiels dominants.

Selon Ritsma, R., 1962 et 1967.

Les intervalles temporels 1-1' ; 1-2' ; 1-3' ; et 2-1' ; 2-2' ; 2-3' et ainsi de suite, fournissent différentes périodes parmi lesquelles le meilleur candidat à la hauteur fondamentale se trouve être celle du « résidu » ou fréquence fondamentale perçue.

Si le son est harmonique (figure 3.29, partie supérieure), cette période et celle du battement coïncident. Si le son est inharmonique (figure 3.29, partie inférieure), différentes périodes sont détectées ; elles correspondent aux différentes hauteurs rapportées par les sujets. Pour l'un et l'autre type d'explication, les sons différentiels (voir chapitre 8, § 2.3) contribueraient à renforcer la perception d'une fondamentale.

Aucun de ces deux groupes de théories ne permet d'expliquer tous les cas de perception de hauteur des sons complexes. Les auteurs s'affrontent en voulant expliquer la perception de la hauteur exclusivement par l'une ou l'autre des modalités de codage. Aujourd'hui des modèles mixtes sont proposés (Moore, 2013, page 222). Nous pensons également qu'il y a lieu de prendre en compte les deux types d'informations (voir aussi Cheveigné, 2004, page 33), sans compter les données cognitives, prédominantes dans ce que nous appelons la zone d'écoute dominante (voir chapitre 6, § 3.5.1).

2.9.4. La hauteur tonale et la hauteur spectrale

L'extension de la psychoacoustique aux sons complexes et en particulier les notions de zone de dominance et de région d'existence nécessitent de prendre en compte la dimension spectrale de la perception de la hauteur, c'est-à-dire la répartition d'énergie dans le spectre du son complexe. Dans une perspective « écologique », Patterson (2010) s'appuie sur l'importance fondamentale des communications vocales interespèces pour proposer une interprétation originale du codage des signaux sonores complexes. Les expériences sont faites avec des sons complexes impulsionnels qui permettent de contrôler de façon indépendante la fréquence fondamentale et le contenu énergétique du son, et rendent compte intégralement des deux aspects, temporel et spectral, de tout signal sonore. Les recherches de Patterson et de ses collaborateurs ont déjà permis de repérer, dans les aires temporales du cortex, des zones distinctes de traitement pour la hauteur fondamentale de l'impulsion et pour la zone spectrale concernée.

Dans la suite de cet ouvrage, et en particulier dans le chapitre 6, nous développons également une double conception de la sensation de hauteur : la hauteur tonale et la hauteur spectrale. La hauteur tonale correspond à cette sensation précise et bien définie que produisent les sons périodiques, comme ceux des instruments de musique entretenus et la voix humaine. La sensation de hauteur spectrale, qui se manifeste principalement lors de variations temporelles du spectre, joue un rôle majeur dans la perception de la parole, mais aussi dans l'interprétation de hauteur de sons musicaux possédant des formants spectraux, ainsi que pour les instruments de percussion dits de hauteur indéterminée (voir chapitre 7, § 2.1.2).

Dans le cadre des connaissances actuelles sur le système auditif, on peut avancer que les théories du codage tonotopique rendent bien compte de la hauteur spectrale, mais que la hauteur tonale serait plutôt codée temporellement.

3. Observations sur quelques phénomènes temporels

Si l'aspect temporel de la perception est abordé en fin de chapitre, c'est que le temps n'est pas une dimension comme les autres. Les questions qu'il soulève traversent tous les champs des connaissances humaines, de la philosophie antique à la physique quantique et, de façon très privilégiée, la création musicale[25].

En ce qui concerne la perception sonore, nous sommes confrontés au fait que le son et l'acte de sa perception se déroulent simultanément, dans le même intervalle de temps. Le problème a été longuement analysé par Husserl[26] qui, pour décrire le déroulement temporel de la perception, s'appuie à plusieurs reprises sur des exemples musicaux. Dans le cadre de ce chapitre de psychophysique, il s'agit d'explorer les seuils temporels du système auditif périphérique, en usant de signaux élémentaires, ne sollicitant pas une interprétation cognitive élaborée.

3.1. Du successif au continu : une frontière physiologique

3.1.1. De la sensation rythmique à la sensation mélodique : une expérience d'écoute

Son 3.17 (33'')

Des clics au son de hauteur définie

À l'écoute d'une succession régulière de clics dont la cadence, d'abord lente, augmente progressivement (Son **3.17**), nous percevons distinctement le resserrement des clics avec l'accélération du mouvement, puis, à un certain moment – variable selon les auditeurs –, un son grave ascendant commence à émerger. Pendant quelques instants, la perception flotte dans une zone mixte, à l'intérieur de laquelle les clics et le son ascendant coexistent. Puis les clics s'estompent et finalement disparaissent, laissant la place au seul glissando ascendant continu. Nous prenons alors conscience que nous avons basculé d'une perception de nature discontinue, de type rythmique, vers une perception de nature continue et mélodique.

La figure 3.31 représente une échelle temporelle pseudo-logarithmique graduée de 1000 à 0,25 ms. Nous y avons mentionné, de part et d'autre de la zone dite mixte, les signes représentant les durées musicales (tempo d'une seconde à la noire) et la position des notes de musique correspondant aux durées de la période (ou à la fréquence).

Il faut souligner que le signal acoustique du Son **3.17** n'a pas changé de nature : il est toujours constitué d'une succession de clics. Seul l'intervalle temporel Δt entre deux clics varie. Pendant la durée de cet exemple l'intervalle passe de 1 seconde à 50 microsecondes (0,05 ms), ce qui correspond au domaine de fréquences allant de 1Hz à 20 kHz. Tant que Δt reste en deçà d'une certaine valeur, les clics successifs sont donc perçus comme des événements isolés dont on peut apprécier le degré de resserrement. Nous restons dans le domaine du rythme. Au-delà, pour de plus faibles valeurs de Δt, les événements isolés ne sont plus séparables, ils fusionnent. L'auditeur perçoit alors une « hauteur » ascendante. Entre ces deux zones très distinctes s'étend une zone mixte plus ou moins bien définie.

25. Pour une revue des écrits relatifs aux rapports entre temps et musique, se reporter à l'ouvrage d'Eric Emery (1998), *Temps et musique*.
26. Husserl, 2002, *Leçons pour une phénoménologie de la conscience intime du temps.*

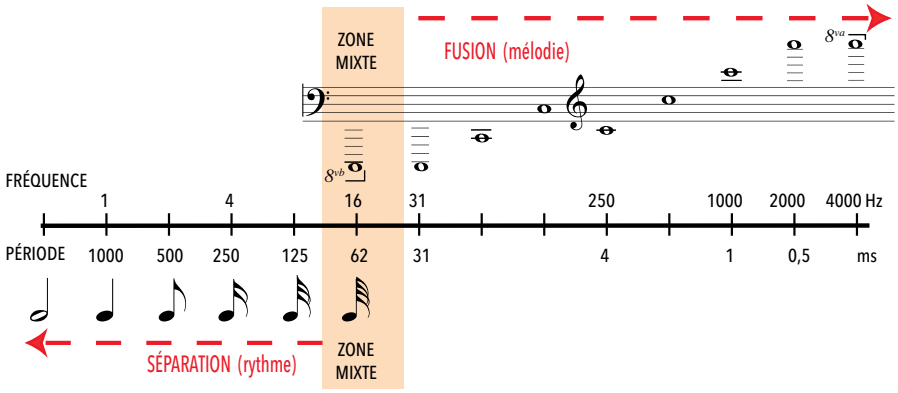


Figure 3.31 Échelle des intervalles temporels. De part et d'autre d'une zone mixte, correspondant à une perception ambiguë, s'étendent vers la gauche le domaine de la perception des rythmes, et vers la droite celui de la perception des fréquences et de la mélodie.

3.1.2. La zone mixte de transition et la limite inférieure de sensation mélodique

L'écoute du Son **3.17**, au cours duquel la perception bascule de la discrimination entre deux impulsions à la fusion mélodique pose le problème d'une *constante de temps d'intégration* qui n'est pas sans rappeler le compromis Δf.Δt = constante. Une bonne analogie visuelle du phénomène est présentée sur la figure 3.32, où l'exemple sonore est analysé avec deux tailles de FFT différentes.

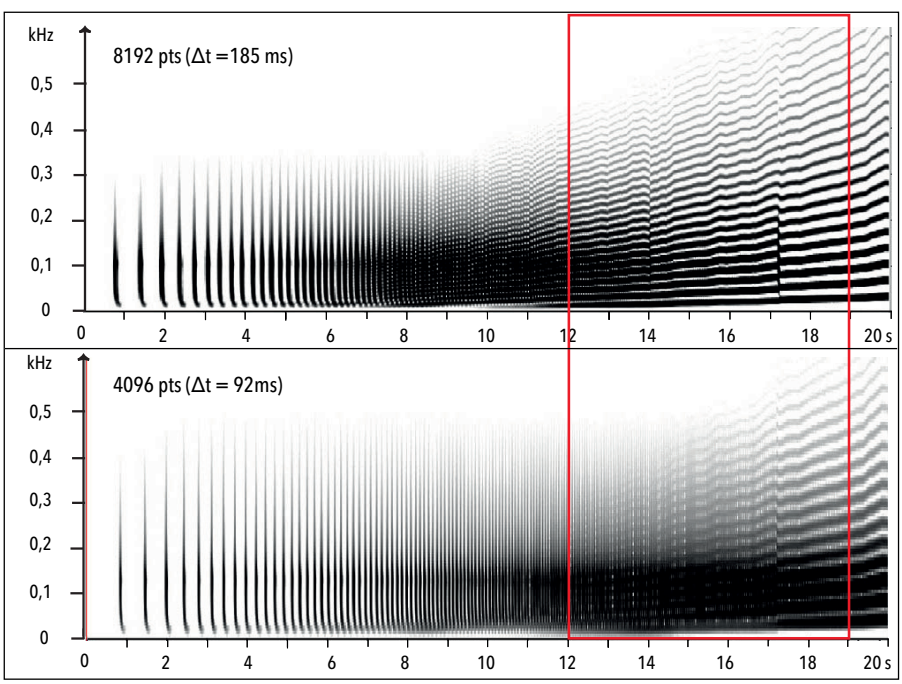


Figure 3.32 Impulsions de fréquence continûment croissante. Analyse spectrale des vingt premières secondes du Son 3.17, avec deux FFT différentes (fréquence d'échantillonnage = 44,1 kHz). Le cadre rouge délimite la zone correspondant aux intervalles de temps compris entre 62 et 30 ms (ce qui correspond respectivement à 16 et 33 Hz).

Au cours de l'accélération, nous avons repéré que l'intervalle entre deux impulsions successives est de 62 ms vers la 12e seconde et 30 ms vers la 19e seconde. Le cadre rouge de la figure délimite à peu près la zone à l'intérieur de laquelle on peut observer, chez la plupart des auditeurs, la transition du discontinu au continu, qui correspond à l'intervalle de fréquence 16-33 Hz[27]. La fréquence la plus basse, 16 Hz, est celle du do_{-1}, limite inférieure de la tessiture des instruments de musique : c'est la note la plus grave d'un jeu d'orgue de 32 pieds[28] qui, en réalité, n'est employée qu'en association avec d'autres, à l'octave supérieure.

Dans une série d'expériences rigoureuses menées en contexte musical, avec des sons impulsionnels, D. Pressnitzer *& al.* ont établi que la limite inférieure de discrimination mélodique était de 31 Hz (do_0). À partir de cette fréquence, la sensation de hauteur est nettement constituée et il est possible de discriminer un intervalle d'un demi-ton. Il s'agit bien d'une limite temporelle de détection, car le filtrage du signal n'affecte pas la valeur trouvée, du moins tant que la fréquence de coupure reste inférieure à 800 Hz : au-delà de cette valeur, la sensation de hauteur s'affaiblit considérablement. Cette expérience a été faite avec trois sujets et il serait intéressant de l'étendre à un groupe de musiciens professionnels pratiquant des instruments de tessiture grave : joueurs de tuba, de contrebasse, de basson et contrebasson. Les explorations que nous avons faites avec les musiciens du Conservatoire de Paris nous ont montré qu'une dispersion manifeste existait. Certains auditeurs perçoivent clairement les intervalles dès que Δt entre deux impulsions avoisine 45 ms (soit 22 Hz) alors que d'autres ne réagissent qu'aux alentours de 20 ms (soit 50 Hz). L'expérience serait donc à reprendre avec un grand nombre de musiciens.

On retrouve des frontières similaires entre discrimination et fusion dans tous les cas où une variation périodique est appliquée à un son. La zone temporelle de transition varie avec la nature des signaux, avec l'enveloppe de la variation et avec l'interprétation cognitive mise en jeu.

3.2. Les variations périodiques de sons de hauteur définie : modulations et battements

3.2.1. Introduction

Les phénomènes de modulation et de battement donnent lieu à une diversité d'effets musicaux dont l'interprétation dépend beaucoup du contexte musical et, d'une façon générale, de la connaissance des canons esthétiques d'une musique donnée. Nous nous limiterons ici à l'analyse de la perception des variations temporelles de sons purs.

Son 3.18 (31'')

Modulation d'amplitude de rapidité croissante

À titre d'introduction, écoutons une expérience similaire à celle des clics, mais appliquée cette fois à un son pur de 525 Hz, dont l'amplitude est modulée avec une cadence qui s'accélère graduellement (Son **3.18**). À l'écoute, nous découvrons à nouveau une frontière entre le suivi temporel de la modulation – tant que celle-ci reste lente – et la sensation d'une hauteur mélodique dès que la modulation est suffisamment rapide. Mais une surprise nous attend : nous n'entendons plus une seule note mais un accord de trois sons.

27. À rapprocher de la vision, pour laquelle la sensation d'un mouvement continu apparaît aux alentours de 25 images par seconde.
28. La hauteur sonore d'un tuyau d'orgue s'exprime en pieds. Le plus grand tuyau d'un jeu de 32 pieds ouvert de la famille des flûtes (par exemple la Montre du Grand-Orgue de Notre-Dame de Paris), mesure 10 mètres de haut. Il émet un son, dont la note fondamentale, do_{-1}, est d'environ 16 Hz.

Des phénomènes analogues se produisent lors de la modulation de fréquence (Son **3.19**). L'analyse spectrographique de ces deux modulations est présentée sur la figure 3.33.

Son 3.19 (31″)

Modulation de fréquence de rapidité croissante

Pour les deux modulations, trois étapes sont intéressantes à détailler comme précédemment : celle des variations lentes qui permettent de suivre la modulation (A), celle des variations rapides donnant lieu à un accord de plusieurs sons (C) et, entre les deux, une zone de transition (B).

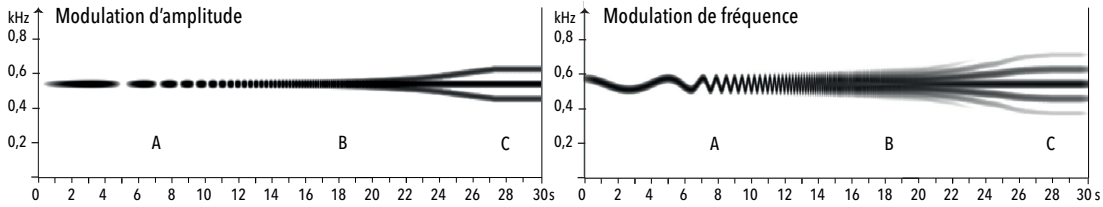


Figure 3.33 Analyse spectrographique de la modulation d'un son pur (525 Hz) par une fréquence croissant graduellement de 0,5 à 84 Hz. À gauche : modulation d'amplitude (Son 3.18). À droite : modulation de fréquence (Son 3.19). Les deux analyses mettent en évidence la transition (B) entre les modulations lentes (zone A) et les modulations rapides (zone C), qui donnent lieu à l'apparition de fréquences collatérales, situées de part et d'autre de la fréquence du son modulé. Ce type de représentation varie avec les paramètres de la FFT (fe = 44,1 kHz ; fenêtre : 4 096 pts et 0,092 s).

3.2.2. Les modulations lentes, inférieures à 12 Hz : suivi de la modulation

Les variations lentes de l'amplitude et de la fréquence sont des composantes fondamentales de l'expressivité musicale. Lorsqu'elles sont périodiques, elles prennent le nom de vibrato.

En psychoacoustique, les modulations sont utilisées pour évaluer les seuils différentiels d'intensité et de fréquence (voir page 104) et pour estimer la largeur des bandes critiques (voir page 99). La fréquence de modulation couramment employée dans les expériences est de 4 Hz. C'est celle pour laquelle nous avons la plus grande sensibilité au taux de modulation, qu'il s'agisse d'une modulation d'amplitude ou de fréquence (voir figure 3.34, page 122). Il en est de même des battements entre deux sons, dont l'effet perceptif est dominé par la modulation d'amplitude. Ainsi, on retrouve cette même valeur de 4 Hz en facture instrumentale pour le vibrato d'amplitude du vibraphone, du tremblant de l'orgue, ou pour l'ondulation de la Voix céleste à l'orgue (voir chapitre 6, page 266).

Quand la rapidité (ou fréquence) de modulation augmente, la sensibilité aux variations évolue différemment pour l'amplitude et pour la fréquence. Le test du Son **3.20** fait entendre un son de 525 Hz modulé en fréquence successivement à 4, 6, 8 et 12 Hz.

Son 3.20 (12″)

Effets de la variation de vitesse de la modulation en fréquence

On se rend compte combien la perception se transforme rapidement. À 4 Hz, le son paraît affligé d'un pleurage important ; à 6 Hz, il scintille dans un intervalle qui paraît plus réduit ; à 8 Hz, l'intervalle de variation semble presque trop petit et la modulation trop rapide et à 12 Hz le son grelotte ! Selon Moore *(op. cit.,* page 197), le mécanisme physiologique de suivi de la modulation de fréquence réagit plus lentement que celui de la modulation d'amplitude, ce qui expliquerait que nous soyons capables, dès 6 Hz, d'abandonner le suivi de modulation de fréquence pour

fusionner les variations en une hauteur moyenne, stable. Simultanément, l'appréciation de la modulation de fréquence change de nature : elle devient une qualité timbrale (voir chapitre 6, § 4). Les mêmes variations de la vitesse de modulation appliquées à l'amplitude paraissent produire des effets plus graduels : les « secousses » s'accélèrent de façon assez monotone. Il faut ajouter que les effets que nous venons de décrire changent considérablement avec la forme de l'amplitude de modulation, et en particulier avec le fait que celle-ci passe ou non par 0.

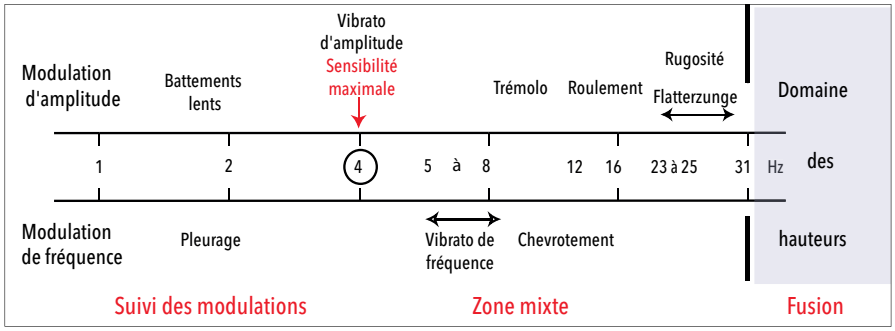


Figure 3.34 Changements perceptifs corrélés à la variation de rapidité (ou fréquence) de modulation. Comparaisons entre la modulation d'amplitude (partie supérieure) et la modulation de fréquence (partie inférieure).

Lorsque la modulation atteint et dépasse 8 Hz, l'effet ressenti par les modulations se déprécie rapidement. En musique, le vibrato d'amplitude devient un *trémolo* et le vibrato de fréquence un *chevrotement*. La combinaison des deux produit le *bêlement*. Notons que les variations périodiques d'amplitude ou de fréquence sont réalisées en musique instrumentale par des mouvements volontaires : celui de la main (jeu du violon), celui du diaphragme (jeu des instruments à vent), celui du larynx[29] (vibrato vocal). Du point de vue physiologique, de tels mouvements sont limités dans leur rapidité. Au-delà d'une fréquence d'oscillation d'environ 8 Hz, les muscles activés se tétanisent (il ne peuvent plus être soumis à contraction).

3.2.3. Les modulations d'amplitude, comprises entre 12 et 30 Hz : la rugosité

Aux environs de 10 à 12 Hz, il n'est plus possible de suivre la modulation (de compter les coups). Nous sommes seulement sensibles aux changements de rapidité en termes de variations de qualité sonore.

L'effet produit par une modulation rapide de l'enveloppe d'amplitude (modulation ou battement) porte le nom de *rugosité*[30]. La rugosité est un des paramètres de l'évaluation psychoacoustique de la qualité des sons générés par les systèmes mécaniques, comme les moteurs et les machines tournantes. Elle a donné lieu à des études systématiques[31] et à la définition d'une unité : l'*asper*[32]. Il s'agit d'une sensation complexe dont l'effet perceptif dépend de la tessiture (fréquence porteuse), de l'amplitude de modulation et de la forme de la modulation. Lorsque celle-ci est très abrupte, ou lorsque des interruptions se produisent, l'effet de rugosité peut devenir

29. Très précisément, c'est la contraction rythmée du muscle crico-thyroïdien qui produit le vibrato vocal (voir Hirano, 1995).
30. Remarquons avec Pressnitzer (1998) que l'équivalent français du terme allemand *Rauhigkeit* serait plutôt *raucité*.
31. Voir Zwicker (1999), Pressnitzer (1998).
32. Un son de 1 kHz de 60 dB, modulé à 70 Hz, au taux de 100 %, a une rugosité de 1 *asper* (Zwicker & Fastl, 1999, § 11.1).

très agressif. Le sifflet à roulette est un bon exemple (Son **3.21**), qui combine une tessiture centrée sur le maximum de sensibilité de l'oreille – autour de 2600 Hz dans notre exemple – et une modulation de fréquence qui varie de 10 à 40 Hz selon le souffle. On ne peut y échapper !

Son 3.21 (4'')

Sifflet à roulette

On rencontre en musique un grand nombre d'effets que l'on peut qualifier de modulations d'amplitude. Certains sont dus, comme le *flatterzunge* de la flûte, à une interruption périodique du son par la langue vibrant dans le débit d'air. La fréquence est de l'ordre de 22 Hz. D'autres, baptisés roulements, sont dus à des instabilités de la production du son, et se rencontrent fréquemment dans les multiphoniques. Il faut aussi considérer la catégorie des *trémolos*, les sons produits par des mouvements d'aller et retour (triangle), de torsion (torsion de l'avant-bras du trémolo d'archet ; torsion du poignet dans le secouement du tambour de basque). Dans ce cas, la fréquence de modulation, voisine de 16 Hz, est double de celle de la limite de contraction d'un seul muscle. Enfin, grâce au mouvement alterné des deux mains, les *roulements* effectués en percussion dépassent 16 Hz sur les surfaces dures (bongos) et peuvent atteindre 20 Hz lorsque l'instrumentiste bénéficie du rebond de la peau (tambour militaire, caisse claire, timbale)[33].

3.2.4. Les modulations rapides : domaine des hauteurs (> 25 Hz)

Une modulation rapide fait basculer dans le domaine des hauteurs, pour une fréquence de transition qui dépend, comme nous venons de le voir, du type de modulation, de l'enveloppe temporelle et de la tessiture. Lors d'une analyse spectrographique, la transition dépend également des paramètres de l'analyse. Dans l'exemple présenté sur la figure 3.33 (page 121), la modulation finale du son de 525 Hz à 82 Hz produit clairement un accord musical constitué des collatérales dont les fréquences sont faciles à trouver. Ce sont 607 Hz (525 + 82) et 443 Hz (525 – 82). La modulation de fréquence possède deux collatérales supplémentaires à 689 et 361 Hz. Le Son **3.22** permet de comparer l'accord résultant de la modulation d'amplitude et celui de la modulation de fréquence. L'accord formé des trois composantes principales[34] est identique, mais les intensités relatives des collatérales sont différentes, ce qui produit, selon les auditeurs, soit un changement de timbre perçu, soit une oscillation dans la prédominance perceptive des composantes.

Son 3.22 (10'')

Modulations d'amplitude et de fréquence à 82 Hz (se reporter aux Sons 3.18 et 3.19)

3.2.5. Les battements, les modulations d'amplitude, les sons différentiels

Les battements se produisent lorsqu'on émet simultanément deux sons de fréquences voisines, par exemple $f1$ et $f2$. En se combinant, les deux ondes primaires produisent une onde résultante dont l'amplitude globale a pour fréquence $f1$-$f2$ (voir Glossaire). En quoi consiste la différence entre la modulation d'amplitude d'un son, et un battement entre deux fréquences ?

Écoutons le Son **3.23**. Nous entendons successivement deux sons purs dont la hauteur, voisine d'un do_4 (525 Hz), est animée d'une variation régulière de l'amplitude, à une cadence de 4/s. Or, le premier est produit par la modulation d'amplitude à 4 Hz d'un son de 525 Hz, et le second par un battement entre deux sons simultanés dont les fréquences sont 523 et 527 Hz. Les deux sons produisent sensiblement le même effet perceptif, du moins tant qu'il s'agit de sons purs.

Son 3.23 (7'')

Son pur modulé en amplitude, puis battement

33. Voir François, 1991, *Percussion et musique contemporaine*, page 165.
34. Dans cet exemple les « notes » des trois composantes forment un accord de quinte diminuée : (la_3) ; (do_4) ; (mib_4).

Son 3.24 (7″)

Son complexe modulé en amplitude puis battements (comparer avec 3.23)

Il en va autrement avec des sons riches en harmoniques, comme ceux des instruments de musique. À titre d'exemple voici la même expérience effectuée avec deux ondes dites en dents de scie (Son **3.24** et figure 3.35). Les effets de la modulation d'amplitude et du battement sont très différents. De fait, dans la modulation d'amplitude, les harmoniques sont modulés en bloc à la même fréquence alors que, pour le battement, ils sont modulés à des fréquences différentes. Toutefois, c'est la fréquence de modulation globale de l'enveloppe qui représente la fréquence du battement.

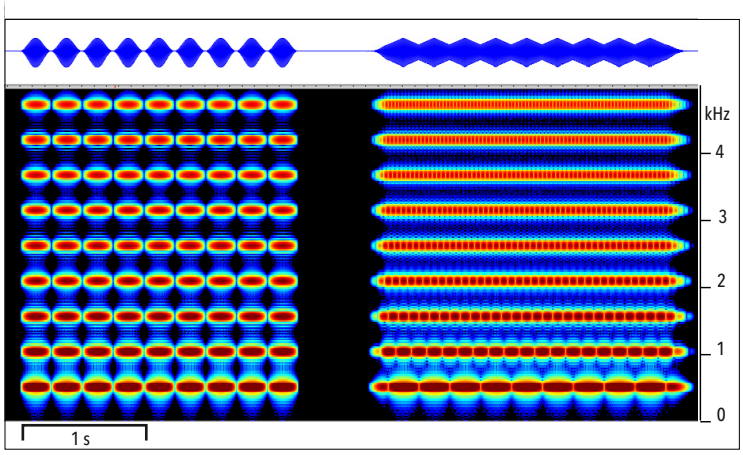


Figure 3.35 Analyse spectrale de deux variations temporelles de l'amplitude sur des sons complexes. À gauche, modulation d'amplitude ; à droite battement.

Helmholtz s'attarde longuement sur la perception des effets dus aux battements, car ils jouent un rôle important dans ce qu'il appelle la « consonance » musicale et dans l'accordage des instruments (1874, page 230).

Son 3.25 (14″)

Battements et tessiture (4 Hz et 8 Hz)

Son 3.26 (12″)

Battements et tessiture (30 Hz et 60 Hz)

Reprenons quelques expériences avec des sons purs, certes peu réalistes, mais très pédagogiques pour notre propos, afin d'explorer ce qui se passe lorsqu'on réalise la même fréquence de battements à des octaves différentes. Prenons par exemple les sons la_2, 220 Hz ; la_3, 440 Hz ; et la_4, 880 Hz. Dans le Son **3.25**, on entend des battements de 4 Hz puis de 8 Hz. La perception du battement ne paraît pas affectée par le changement de fondamentale. Passons à 30 Hz et 60 Hz (Son **3.26**). Curieusement, sur la note grave (la_2), nous entendons un intervalle (respectivement un ton et une quarte) alors que sur la note aiguë (la_4) apparaît seulement la sensation d'une rugosité plus ou moins serrée.

Ces deux expériences démontrent que la fréquence critique de basculement du suivi de la modulation à la perception de fréquences autonomes dépend de la tessiture, donc des fréquences fondamentales des deux sons du battement.

Le changement de perception des battements, et d'une façon plus générale des modulations d'amplitude, avec la tessiture, est à mettre en relation avec l'existence des bandes critiques. On considère que le passage de la sensation de rugosité à celle des deux fréquences séparées correspond justement à la largeur d'une bande critique.

Si l'on se reporte à la figure 3.19, page 100, on lit qu'à 200 Hz la largeur d'une bande critique est d'environ 50 Hz, et à 900 Hz, elle est d'environ 115 Hz. Autrement dit, les deux sons de 820 et 880 Hz tombent dans la même bande critique et donnent lieu à la modulation suivie de l'amplitude. Au contraire, 220 et 160 Hz affectent deux bandes critiques distinctes et sont donc perçus comme des fréquences distinctes. Cette interprétation est un bon guide pour la perception, mais ne saurait être prise à la lettre. En effet, plusieurs auditeurs entendent clairement un intervalle d'un ton entre 220 et 190 Hz, alors que l'écart entre les deux sons est inférieur à la largeur théorique de la bande critique. Ajoutons que lorsque la fréquence des battements est très grande et que l'amplitude globale est importante, des sons différentiels apparaissent, surtout s'il s'agit de sons purs.

3.3. La notion de discrimination temporelle

Jusqu'ici nous avons examiné les changements de la perception d'une succession d'événements sonores en fonction de la rapidité de succession. D'autres expériences s'attachent à caractériser nos capacités à détecter la plus petite discontinuité perceptible entre deux événements successifs, ou une brève interruption dans un son continu.

3.3.1. Le TFT d'Émile Leipp

Une des lignes directrices de la recherche développée au LAM par Émile Leipp a été de mettre l'accent sur l'importance des variations temporelles dans la perception des sons réels de la musique[35]. Il a imaginé un petit appareil portable pour tester l'oreille, le TFT (test fréquence temps), comprenant un générateur de fréquences pour explorer la limite supérieure d'audibilité, et un générateur délivrant une salve de 5 clics (Leipp, 1977). L'appareil permet de régler l'intervalle entre 2 clics, dans une gamme de valeurs allant de 1 à 200 ms. Nous avons ainsi pu explorer les réactions de nombreux visiteurs du laboratoire. L'intervalle discriminé par la moyenne des personnes se situait entre 5 et 10 ms. Pour quelques rares sujets, il fallait aller jusqu'à 25 ms. La majorité des musiciens instrumentistes se situait vers les plus petites valeurs, 1 ou 2 ms. L'expérience étant effectuée avec un protocole invariable, ces résultats ont une valeur comparative : les musiciens ont montré une meilleure discrimination temporelle. Mais de quelle discrimination s'agit-il ?

3.3.2. Les difficultés inhérentes à l'exploration temporelle

Lorsqu'on introduit un double-clic dans une séquence, celui-ci peut être discriminé par une différence d'intensité – un double-clic paraît plus intense qu'un clic simple – ou par un changement de hauteur, car le double-clic a une hauteur spectrale différente. Autrement dit, la discrimination ne se fait pas directement sur le paramètre temporel que l'on souhaite explorer, mais sur un effet induit de variation d'intensité ou de hauteur. Pour contourner ces difficultés, une technique consiste à présenter des paires de signaux dont l'une est le renversement temporel de l'autre, ce qui a pour effet d'égaliser l'intensité et le spectre. Une autre technique, qui concerne les expériences dans lesquelles on doit détecter des interruptions dans les signaux, consiste à ajouter un bruit blanc masquant les phénomènes perturbateurs, comme l'élargissement de spectre se produisant à la transition signal/silence. On trouvera un exposé détaillé de ces expériences dans l'ouvrage de Moore, 2013 (chapitre 5, p. 163). D'une façon générale, les perturbations d'amplitude, les « trous » opérés dans des bruits ou dans des sons purs sont discriminés pour des valeurs comprises entre 1 et 5 ms, selon le type de signal.

35. Variations négligées dans les recherches des années 1960 (à l'exception de Winckel, 1960).

La discrimination des modulations d'amplitude, c'est-à-dire la détection de la plus faible valeur du taux de modulation, produit des résultats beaucoup plus complexes. On retrouve trois zones différentes : une zone des modulations lentes dans laquelle le seuil de discrimination de la modulation est constant, quelle que soit la fréquence de modulation – c'est une discrimination temporelle –, une zone de transition dans laquelle le seuil augmente, et une zone de modulations rapides dans laquelle interviennent les changements de qualité sonore dus aux raies collatérales, et où la discrimination est donc spectrale.

3.3.3. La perception des décalages entre composantes d'un son complexe : les transitoires

La détection d'asynchronies entre composantes simultanées a aussi fait l'objet de plusieurs expériences. Avec des sons de synthèse, on peut réaliser des décalages très précis au début ou à la fin du son (figure 3.36). Les résultats montrent que les décalages sont mieux discriminés avec un son périodique qu'avec un son apériodique, et mieux perçus avec un son musical qu'avec un son vocalique.

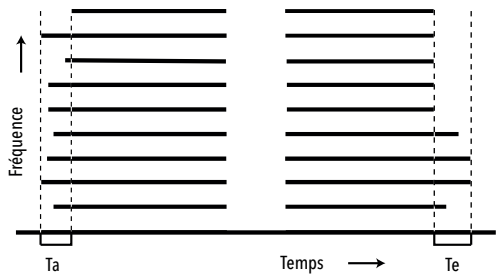


Figure 3.36 Asynchronie à l'attaque (gauche) et à l'extinction (droite).

Toutes les expériences montrent que l'on discrimine beaucoup mieux le décalage initial ou transitoire d'attaque, pour lequel les valeurs sont inférieures à 1 ms, que le décalage final ou transitoire d'extinction, pour lequel les valeurs sont de 3 à 30 ms selon les conditions (Moore, page 190).

On sait que les transitoires jouent un rôle important dans la perception de la qualité, voire de l'identité des sons. Cependant, ils se modifient au cours de la propagation des ondes, ils sont affectés par les mouvements de la tête, et changent d'une fois à l'autre pour un instrument donné. Mieux vaut, semble-t-il, ignorer certaines de ces variations, pour préserver l'identité primordiale du son. Finalement, la question de savoir si les transitoires doivent être discriminés en tant que tels, ou au contraire être intégrés à l'ensemble du signal dont ils modifient la qualité générale, se pose dans le cadre de la perception du timbre (voir le chapitre 7, § 2.3.4, page 311).

3.4. Résolution temporelle ou intégration temporelle ?

Les résultats des expériences de perception temporelle ont conduit les chercheurs en psychoacoustique à introduire dans leur modèle perceptif plusieurs étapes de traitement, dont l'une rend compte des capacités de discrimination d'intervalles à petite échelle, et l'autre réalise une opération de lissage, d'intégration à une échelle temporelle plus grande.

À plusieurs reprises, nous avons constaté dans nos expériences un basculement entre le suivi perceptif d'une succession d'événements lents et la fusion de ces mêmes événements dans une perception de nature différente, dès que l'intervalle de succession dépasse une valeur seuil. Le mécanisme intégrateur qui donne naissance à la sensation de hauteur mélodique est sans doute purement physiologique. Il entre en action dès que la succession des clics (voir § 3.1.2) dépasse 30 Hz, ce qui correspond à environ 60 ms. Cette valeur varie selon les sujets et serait donc caractéristique d'une personne donnée. Dans le cas des modulations d'amplitude ou de fréquence, les seuils d'intégration sont moins bien définis. Ils dépendent, comme on l'a vu, de la fréquence porteuse et comportent sans doute une composante cognitive relevant de l'apprentissage. Le fait de pouvoir fusionner les variations de fréquence d'un vibrato en une hauteur unique, dès 6 Hz, semble plus lié à un acquis culturel qu'à une caractéristique physiologique.

Les aptitudes à la musique requièrent la conjugaison simultanée de bonnes capacités de discrimination temporelle et d'intégration à différentes échelles. L'intégration donne accès à la sensation de hauteur ; la discrimination permet d'apprécier l'infinie variété des petites différences qui se produisent au cours de l'interprétation musicale. Il existe vraisemblablement plusieurs niveaux d'intégration temporelle, qu'un auditeur active alternativement selon les musiques, selon ses motivations et les modes d'écoute qu'il met en œuvre. Et, pour certains (improvisateurs, chefs d'orchestre, interprètes bilingues), le traitement temporel est nécessairement *polyphonique*.

4. La localisation des sons et la perception de l'espace

4.1. La localisation et la latéralisation auditive

La capacité à localiser les sons dans l'espace fait partie des aptitudes fondamentales des êtres vivants qui doivent pouvoir repérer la position d'une proie[36] ou éviter un prédateur en toutes circonstances. En l'absence de données visuelles, le son permet, sur la seule information de l'intensité, de savoir si la source vient de la gauche ou de la droite, si elle est proche ou lointaine et, grâce aux caractéristiques de directionnalité de l'ensemble pavillon + conduit auditif (voir figure 3.3, page 83), d'estimer avec une seule oreille dans quelle portion du plan latéral elle se trouve. Mais nous avons deux oreilles et, spontanément nous tournons la tête dans la direction du son. Ce réflexe fournit une réponse rapide et précise, car il repose sur une « mise à zéro » des différences entre les signaux captés par les deux oreilles.

Pour résoudre des problèmes tels que la détection de sirènes dans la brume ou le repérage d'avions ennemis pendant la Première Guerre mondiale, les ingénieurs ont imaginé des dispositifs pour assister l'oreille dans sa fonction de pavillon (concentrer l'énergie, détecter l'incidence du son) et pour accroître l'écart entre les oreilles afin de gagner en précision temporelle (figure 3.37, page 128). Ces prothèses lourdes et encombrantes furent rapidement abandonnées au profit du radar, mais plusieurs indices perceptifs avaient ainsi pu être expérimentés.

36. Voir la vidéo d'un renard chassant dans la neige, sur le site Internet de Jan Schnupp : auditoryneuroscience.com/foxInSnow.

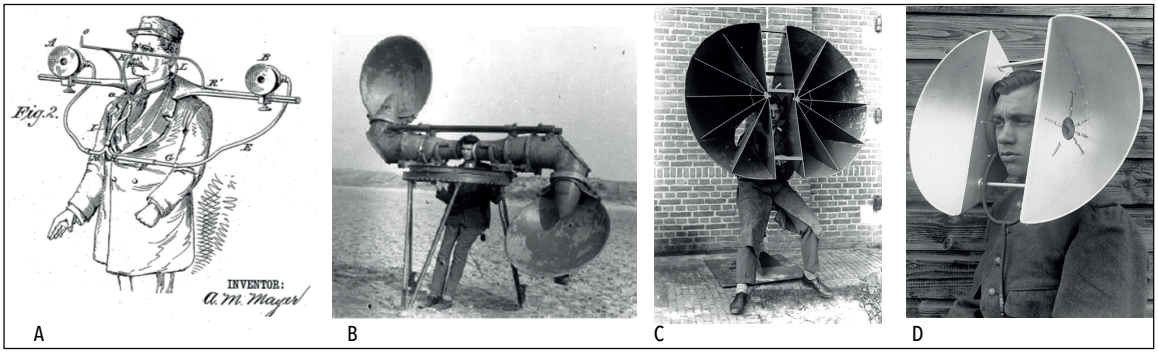


Figure 3.37 Quelques « prothèses » d'oreilles pour affiner la localisation des sons dans l'espace. A : le topophone de Mayer (brevet 1880), développé pour aider les navigateurs à repérer les sirènes de brume. B à D : photos de dispositifs développés pour aider au repérage des bruits aériens pendant la Première Guerre mondiale.

Source des clichés B à D : musée Waalsdorp, Den Haag, Pays-Bas. Voir la documentation sur le site www.museumwaalsdorp.nl/en/airacous.html.

L'étude de la localisation auditive s'attache à identifier les paramètres pertinents sur lesquels se fonde le système auditif pour estimer la position d'une source dans l'espace, et pour en quantifier les performances, en l'absence de tout mouvement de la tête. Une des premières expériences réalisées avec des sons purs est celle de Stevens & Newman (Stevens, 1938, page 176), effectuée dans des conditions acrobatiques, puisque le sujet était hissé au sommet d'une haute cheminée, sur le toit d'un immeuble.

L'intérêt pour ce domaine de recherches s'est accru avec le développement des nouvelles techniques de reproduction du son, la stéréophonie, et plus récemment la généralisation de l'écoute au casque.

Dans cette nouvelle situation, il devient possible de décorréler totalement les signaux sonores parvenant à chaque oreille, donc d'étudier beaucoup plus précisément les phénomènes neurophysiologiques du traitement des signaux binauraux, et de mieux cerner les paramètres qui permettent à un auditeur de recréer intérieurement un espace acoustique dont il est privé par les écouteurs.

Il faut donc distinguer :

- les recherches sur la *localisation auditive*, qui ont pour objectif de quantifier les indices acoustiques utilisés par un auditeur pour repérer la direction d'un son avec les oreilles libres et la tête immobile ;
- les recherches sur la *latéralisation auditive* d'un auditeur équipé d'un casque d'écoute, qui permettent en quelque sorte de préciser les résultats précédents, puisqu'il devient possible de découpler les différences de temps et d'intensité entre les signaux gauche et droite.

La tête d'un auditeur étant au centre d'une sphère, les trois plans remarquables qui concernent la localisation auditive humaine sont : le plan horizontal passant par l'ouverture des canaux auditifs ; le plan vertical sagittal (ou médian) passant par l'axe de symétrie du squelette, à égale distance des deux oreilles, le plan frontal passant par les deux oreilles à angle droit du précédent (voir figure 3.38).

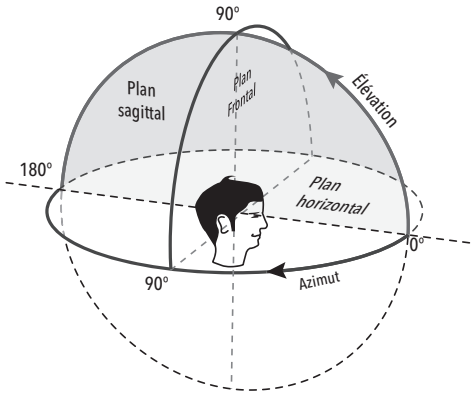


Figure 3.38 Localisation auditive : les trois plans de repérage pour l'étude de la localisation des sons. Les coordonnées angulaires sont l'azimut dans le plan horizontal et l'élévation dans le plan sagittal (ou médian).

4.2. La localisation dans le plan horizontal

« C'est à travers la situation de notre corps que nous saisissons l'espace extérieur », rappelle le philosophe Merleau-Ponty (1962, page 403). Or, les principaux sens qui nous permettent d'accéder à des événements lointains (audition, vision, olfaction) ont leurs capteurs rassemblés à l'avant de la tête, laquelle tourne préférentiellement dans le plan horizontal, dans un espace angulaire moyen de 120°. Par conséquent, les résultats des expériences portant sur la localisation auditive sont particulièrement abondants dans la partie frontale du plan horizontal passant par la tête et on y détecte les meilleures performances. On peut schématiser la situation de la façon suivante (voir figure 3.39).

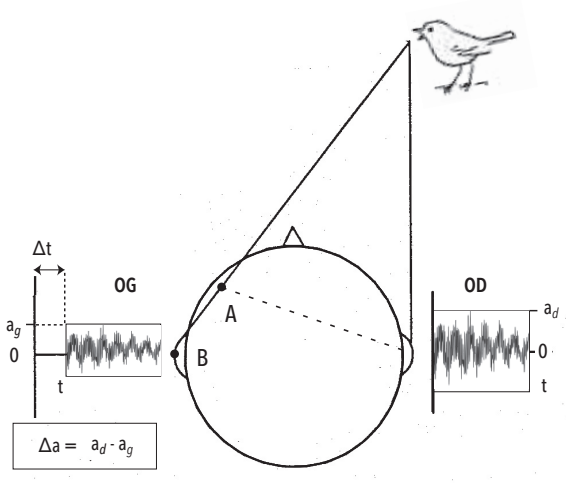


Figure 3.39 Écoute binaurale et représentation des différences perceptibles entre les signaux parvenant à l'oreille droite (plus proche) et à l'oreille gauche.

Un oiseau chante sur la droite de l'auditeur. Figurons par deux droites les ondes qui parviennent aux oreilles droite (OD) et gauche (OG). On remarque que : 1) le son arrive directement à l'OD alors que l'OG est en partie « à l'ombre » de la tête ; 2) le trajet du son vers l'OG est plus long (partie AB) que celui vers l'OD proche de l'oiseau ; 3) l'angle d'incidence du son sur les pavillons change aussi. Les différences entre les signaux sonores parvenant aux deux oreilles sont schématiquement indiquées sur la figure. Ce sont : une différence d'amplitude (Δa), un écart temporel (Δt) entre les signaux droite et gauche, ainsi que des différences dans le contenu spectral du son. Ces données sont traitées dans les centres nerveux où se croisent les voies nerveuses issues de chaque oreille (le noyau cochléaire et l'olive supérieure, voir page 89).

Par cet exemple familier, on voit que la tête a une incidence directe sur la disparité des signaux binauraux : elle est un obstacle à la propagation des ondes sonores du côté opposé à la source, et elle détermine la distance entre les deux oreilles. Précisons quelques ordres de grandeur. La propagation du son fait intervenir la longueur d'onde (λ) et la célérité du son (c) que nous estimerons à 340 m/s. Le diamètre d'une tête moyenne, 21 cm, est la longueur d'onde d'un son de 1600 Hz (f = c/λ, soit 340/0,21 = 1600 Hz). Les fréquences dont la longueur d'onde est plus petite, donc les sons plus aigus que 1600 Hz, seront affectées par cet obstacle. Par ailleurs, l'écart moyen entre les tympans, estimé à 17 cm, est la différence de trajet maximale entre les ondes qui atteignent l'OD et celles qui atteignent l'OG. Le temps de propagation correspondant (0,5 ms), est la période d'un son de 2000 Hz. Les décalages produits par des sons de plus basse fréquence seront aisément détectés. La différence de trajet s'interprète comme différence de phase ($\Delta\phi$) lorsque l'expérience est conduite avec des sons sinusoïdaux, et comme différence de temps (Δt) pour les signaux complexes.

Les différences d'intensité et de temps perçues entre les signaux gauche et droite sont désignées en psychoacoustique par les sigles ILD pour l'intensité (*Interaural Level Difference*), et ITD pour le temps (*Interaural Time Difference*). Lord Rayleigh avait déjà mentionné la complémentarité des deux paramètres dans le cadre de ce qu'il a nommé *duplex theory* (Moore, page 238). En pratique, on admet que l'écart d'intensité est un indice efficace pour les fréquences supérieures à 1500 Hz, alors que l'écart temporel agit plutôt pour les fréquences inférieures à 1500 Hz. Ces données varient selon les individus, mais toutes les expériences montrent que dans la zone de transition, comprise entre 1000 et 2000 Hz, les performances en localisation sonore sont moins bonnes.

4.2.1. Les capacités de discrimination dans le plan horizontal (sons sinusoïdaux)

Soit une source sinusoïdale, émise face à un sujet (azimut 0°) dont la tête est fixe. Quel est le plus petit déplacement angulaire perceptible ? On nomme MAA (*Minimum Audible Angle*) le paramètre mesuré. La figure 3.40 montre que la valeur la plus faible est de 1° pour les fréquences comprises entre 300 et 1 000 Hz (courbe noire, 0°).

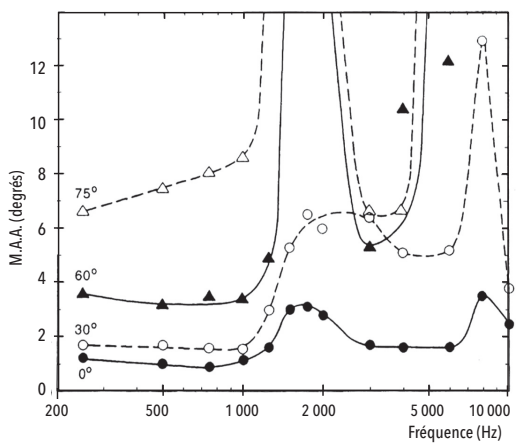


Figure 3.40 Angle minimum détectable dans le plan horizontal ; mesures en fonction de la fréquence, pour quatre incidences de la source : azimut = 0° ; 30° ; 60° ; 75°.
Moore, B., 2013, figure 7.5, (selon Mills, 1958).

Cette valeur augmente dans l'intervalle 1000-3000 Hz, puis à nouveau au-delà de 8 000 Hz. Pour toutes les autres positions d'azimut, les valeurs de MAA sont supérieures à celles de l'azimut 0° (comparez les courbes à 30, 60 et 75°). Remarquez aussi que lorsque la fréquence est comprise entre 1000 et 3000 Hz, seules les positions d'azimut de 0° et 30° permettent une discrimination. Pour toutes les fréquences, les positions 0° (devant) et 180° (derrière, non représentée ici) sont confondues.

4.2.2. Les expériences de latéralisation

Vocabulaire. Les expériences d'écoute au casque permettent de sélectionner les oreilles à exciter – une oreille (écoute *monaurale*) ou les deux simultanément (écoute *binaurale*) – et de contrôler les signaux sonores : même signal aux deux oreilles (condition *diotique*) ou signaux différents à chaque oreille (condition *dichotique*).

Une expérience de latéralisation est souvent conduite de la façon suivante. Un signal binaural diotique est émis en premier comme référence ; il est suivi d'un signal dichotique porteur de différences en temps ou en amplitude. L'auditeur interprète les différences par référence à la situation décrite sur la figure (page 129). Le son est localisé sur l'oreille qui reçoit le son le plus intense ou celui qui arrive en premier. Le plus faible écart temporel détecté dans ce type d'expérience est de 10 µs (0,01 ms) pour la fréquence 500 Hz (Yost, 1987). Les expériences de latéralisation sont nombreuses et complexes, car les résultats dépendent beaucoup des signaux utilisés, mais aussi des sujets. On retiendra qu'un écart temporel de 0,3 ms est un seuil de détection temporel perçu par la majorité des sujets. Certaines personnes sont manifestement plus sensibles aux écarts temporels (ITD) qu'aux écarts d'intensité (ILD) qui doivent dépasser 1dB. Il est intéressant de pratiquer individuellement quelques tests d'écoute avec un très bon casque. Nous recommandons les exemples sonores accompagnant le livre de Jan Schnupp (2011) ainsi que ceux de Duda (1996) publiés dans un CD joint à la publication.

Parmi les phénomènes auditifs spécifiques de l'écoute au casque, signalons aussi les battements binauraux. L'expérience est simple à mettre en œuvre. Enregistrons dans un fichier bi-pistes deux sons sinusoïdaux de même intensité et de fréquence légèrement différente, par exemple 440 Hz à gauche et 444 Hz à droite. Diffusés sur une paire d'enceintes, les deux signaux sonores se combinent et nous entendons une variation de l'amplitude de 4 Hz (voir page 123). Écoutés au casque, les signaux affectent de façon séparée chacune des deux oreilles, et la différence de fréquence, qui correspond aussi à des décalages dans les maxima d'amplitude, produit une oscillation de la latéralisation auditive à la fréquence du battement, ici 4 fois par seconde. Vous trouverez un exemple sonore sur le site Internet de Jan Schnupp & coll. : auditoryneuroscience.com/binauralBeats.

4.3. La localisation dans le plan vertical : plan sagittal (ou médian)

Les tests portant sur la localisation des sources dans le plan vertical sagittal se sont révélés décevants avec des sons purs. En utilisant des bruits de bande étroite, Jenz Blauert (1983) a défini des « bandes directives » liées à certaines fréquences : des sons dont le maximum est à 8 kHz étaient perçus au zénith ; les sons inférieurs à 500 Hz et ceux de 3 kHz étaient plutôt localisés derrière (azimut 180°), alors que ceux de 1 kHz étaient perçus de face (azimut 0°).

4.4. La prise en compte des pavillons

Paralllèlement à ces travaux, assez éloignés d'une situation réelle, tant par la nature des stimuli utilisés que par les conditions d'écoute, des chercheurs se sont interrogés sur la contribution des pavillons de l'oreille à notre aptitude à localiser les sons. Il a été constaté que, lorsqu'on ne considère comme point d'entrée que le trou d'ouverture des conduits auditifs, il existe un cône de confusion à la surface duquel les changements d'incidence des sons purs sont indiscernables (voir Moore, p. 248). Ces résultats ne sont pas valides pour les sons complexes. L'intérêt s'est alors porté sur la géométrie particulière des pavillons d'oreille dont les petites anfractuosités diffractent les ondes sonores de façon spécifique selon l'incidence des sons et selon leur fréquence. Toutefois, ces effets sont dépendants de l'anatomie individuelle : il font partie intégrante de notre expérience personnelle de la sensation spatiale. Certains auditeurs semblent y être plus sensibles que d'autres. L'établissement des courbes de la « fonction de transfert liée à la tête » ou HRTF (*Head Related Transfer Function*) pour une personne donnée, est une opération longue mais qui semble prometteuse puisque, selon Moore[37], leur prise en compte a pour effet que « le son n'est plus perçu quelque part à l'intérieur de la tête, mais semble provenir de l'espace extérieur ».

4.5. L'effet de précédence ou loi du premier front d'onde

L'effet de précédence, appelé aussi effet Haas, désigne le fait que nous assimilons la direction d'une source à celle d'où provient le premier front d'onde. Les signaux dus aux premières réflexions qui arrivent avec un léger décalage temporel sont occultés au profit du premier, même s'ils sont plus intenses, du moins tant que le décalage Δt reste faible et que la différence d'intensité ne dépasse pas 10 dB. Le décalage critique Δt varie de 15 à 30 ms selon le type de signal (impulsionnel ou progressif) et selon les propriétés acoustiques du lieu (amorti ou résonant). Exploité en sonorisation grâce aux lignes à retard, cet effet permet d'amplifier une source sonore sans introduire l'effet, fortement désagréable pour un auditeur, d'une dissociation entre la localisation visuelle d'un orateur et sa localisation auditive.

4.6. L'estimation de la distance

L'estimation de distance d'une source est très peu précise avec des sons sinusoïdaux. En champ libre, pour un son pur, seule l'atténuation donne une indication de distance. Dans un lieu non absorbant, le rapport son direct/son réverbéré est la première indication.

Lorsque le signal est complexe, intervient l'atténuation rapide des fréquences aiguës avec la distance. À l'opéra, on reproduit l'effet d'éloignement en plaçant des instrumentistes dans les coulisses, car les chicanes du décor et les rideaux de scène jouent le rôle de filtre en donnant à l'auditeur l'illusion de la distance. De façon similaire, à l'orgue, le clavier « d'écho », dont les tuyaux sont situés dans le soubassement du buffet principal, produit un son affaibli et appauvri en fréquences aiguës.

4.7. La localisation en situation réelle

Les résultats obtenus avec les expériences que nous venons de relater, en particulier ceux issus des études de latéralisation, ont permis de progresser dans la compréhension des traitements neurophysiologiques sous-jacents à notre interprétation de la spatialisation sonore, mais il est encore difficile de faire le lien avec les situations réelles pour les raisons suivantes.

37. « *The pinae alter the sound in a way that causes the sounds to be perceived as externalized.* » p. 250.

1 – La tête est mobile. Comme nous l'avons évoqué au début de cette section, le fait de « pointer le nez » dans la direction du son est un réflexe rapide d'annulation des différences entre OD et OG, qui ne nécessite pas une estimation des indices mentionnés. En outre, de même que nos yeux sont animés d'incessants petits mouvements, nous bougeons la tête, même imperceptiblement, sans discontinuer. Ce faisant, nous provoquons de micro-changements de direction pour lesquels nous avons une très grande sensibilité différentielle. Il est bien possible que ces phénomènes participent du plaisir de l'écoute naturelle des sources mécaniques, dont les variations de rayonnement sont, pour certains instruments comme le violon, très rapides et très complexes selon la note jouée (voir Weinreich, 1997). Pour ceux d'entre nous qui y sont sensibles, ces effets ne se retrouvent ni à l'écoute sur enceintes, ni à l'écoute au casque, conditions d'écoute que nous acceptons faute de mieux. Les systèmes actuels de recréation d'un champ sonore complexe sont encore expérimentaux[38].

2 – Dans l'écoute habituelle, nous avons affaire à des sources connues ; nous avons engrangé, à leur sujet, un grand nombre de connaissances sur les transformations sonores provoquées par les variations de directions, principalement dans le domaine spectral. Dans le cas de la localisation verticale (sagittale), intervient aussi l'expérience acquise sur le comportement des animaux ou des machines : les sons aigus évoquant le chant d'oiseaux habituellement perchés en hauteur, les bruits graves ceux de machines au sol. Le mouvement éventuel de la source est aussi une indication importante. Des remarques similaires peuvent être faites concernant l'appréciation de distance de sources connues.

3 – Enfin, nous nous trouvons fréquemment dans un lieu plus ou moins clos, environné d'ondes sonores venant de toutes les directions. Il faut certainement beaucoup d'expérience pour intégrer le traitement des ondes et de leurs réflexions, comparer les ILD et les ITD, mais dans une situation ordinaire où nous connaissons les sources sonores, nous pouvons anticiper leur emplacement. Dans l'extrait suivant, Marcel Proust décrit à merveille comment la connaissance du contexte et la vision de l'objet à repérer permettent de déjouer les pièges des réflexions multiples :

> *J'entendais le tic-tac de la montre de Saint-Loup, laquelle ne devait pas être bien loin de moi. Ce tic-tac changeait de place à tout moment, car je ne voyais pas la montre ; il me semblait venir de derrière moi, de devant, d'à droite, d'à gauche, parfois s'éteindre comme s'il était très loin. Tout d'un coup je découvris la montre sur la table. Alors j'entendis le tic-tac en un lieu fixe d'où il ne bougea plus. Je croyais l'entendre à cet endroit-là ; je ne l'y entendais pas, je l'y voyais, les sons n'ont pas de lieu.*
>
> Proust, M., 1994, p. 79.

En situation ordinaire, c'est-à-dire sans casque ni écouteur, nous captons avec chaque oreille deux flux acoustiques similaires, mais différents. Par expérience, nous apprenons qu'ils se réfèrent à une seule source, mais surtout nous pouvons tirer parti des différences entre ces signaux pour en extraire des informations sur la position et la distance de la source sonore. Il reste que l'appréciation des qualités du timbre, qui repose en partie sur le contenu spectral, conduit à penser que dans ce domaine aussi nous oscillons sans cesse entre l'appréciation fine des variations du champ sonore changeant temporellement et différemment aux deux oreilles, et le besoin d'une synthèse dynamique pour estimer le timbre d'un instrument à un moment donné, comme le montre l'analyse présentée dans l'encadré de la page 134.

38. Voir Warusfel, O., 2008, « La spatialisation du son », *Pour la Science,* n° 373 ; Noistermig, M., *& al.,* 2012. Dispositif de spatialisation sonore 3D à l'espace de projection de l'IRCAM – un réseau de 345 haut-parleurs pour une restitution par WFS et HOA. *Acoustique et Techniques*, n° 71.

Variabilités locales du champ sonore et permanence perceptive

Le son que nous entendons en un point donné de l'espace résulte de la combinaison de toutes les ondes sonores issues de la structure vibrante, et de celles qui sont réfléchies par des obstacles situés sur leur trajet. Chaque microphone d'un couple stéréo fournit un signal particulier. Comparons le son d'une note enregistrée simultanément en deux points – proches mais différents – de l'espace sonore.

L'analyse spectrographique de la figure 3.41 montre d'importantes variations spectrales entre les deux signaux : le signal qui parvient à l'oreille gauche est différent de celui de l'oreille droite. Pourtant, nous entendons un seul son avec une qualité globale homogène, de même que les deux images captées par les yeux fusionnent en une seule vue.

Un auditeur humain a la capacité d'extraire, à travers les variabilités de la production sonore et du rayonnement, les indices constitutifs de l'identité de la source et ceux qui sont caractéristiques de sa qualité, et ce, même lorsqu'il se déplace dans le champ sonore. De façon analogue à la permanence visuelle – les objets ne se déforment pas lorsque nous tournons la tête – il s'agit d'un processus cognitif que chacun de nous développe et affine au cours de son expérience d'écoute. Mais si les conditions d'enregistrement et de reproduction du signal sonore n'altèrent pas la reconnaissance des sources (timbre identitaire), elles peuvent avoir une incidence notable sur l'appréciation de leurs qualités (timbre qualitatif).

Son 3.27*a (3'')

Son stéréophonique

Son 3.27*b (6'')

Son monophonique ; voie gauche puis voie droite

Son 3.27*c (3'')

La 1re note : voie gauche, puis voie droite

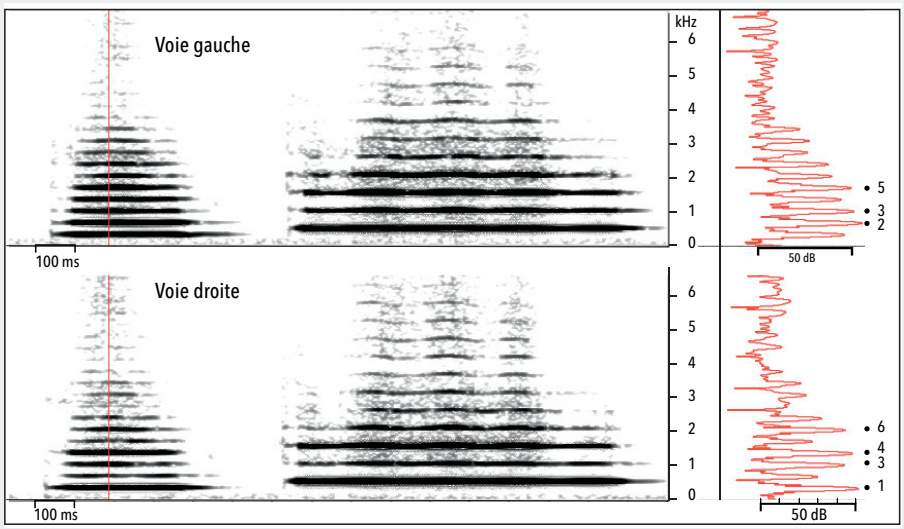


Figure 3.41 Analyse du son d'une flûte traversière enregistré avec les deux microphones directionnels d'un couple stéréo ; en haut, la voie gauche et, en bas, la voie droite. Le spectre de la première note, capté à la position du curseur rouge, est visible à droite du sonagramme. Il met en évidence les importantes différences de qualité sonore que l'on perçoit lorsqu'on écoute cette note en séparant chaque canal (Son 3.27c). Les harmoniques les plus intenses sont, pour le canal gauche : 2, 3, 5 ; pour le canal droit : 1, 3, 4, 6. Ces différences, intégrées par l'auditeur dans la perception globale de la sonorité, contribuent à l'impression d'espace et de volume du son. Elles mettent en lumière l'important problème de la prise de son et son incidence sur les analyses acoustiques du timbre.

L'art des preneurs de son consiste précisément à nous offrir, non une restitution fidèle, inatteignable, mais une recréation de l'espace sonore qui nous permet d'activer l'immensité de nos expériences d'écoutes déjà mémorisées, afin de participer à la recréation de celle qui nous est proposée.

Pour compléter cette présentation sommaire, nous renvoyons le lecteur aux ouvrages cités de Canévet *(Audition binaurale et localisation auditive*, 1989) ; Hugonnet et Walder (*Théorie et pratique de la prise de son stéréophonique*, 1995) ; Brian C. J. Moore (*Psychology of Hearing*, chapitre 7, 2004) ; Jan Schnupp & coll. (*Auditory Neuroscience*, 2011).

5. Les sons du chapitre 3

Important. Pour toutes les expériences de perception sonore et en particulier pour la détection d'intervalles très faibles comme ceux des seuils différentiels il est intéressant de contrarier sans cesse l'impression perceptive que l'on a. Si vous entendez un intervalle ascendant, essayez de l'entendre descendre. Si vous y parvenez, c'est qu'il s'agit d'une suggestion et non d'une perception objective.

Pour l'écoute de ces expériences une bonne qualité de reproduction sonore est nécessaire, en particulier pour les sons purs de basse fréquence (exemples **3.9** et **3.13**) et pour les exemples **3.15** (a-e) et **3.27**. Certains exemples sont à écouter au casque.

Les réponses aux tests se trouvent à la fin de ce paragraphe.

5.1. Simulation de surdités : filtrage de la voix parlée (voix féminine)

Son 3.1 – Voix parlée : filtrage passe-bas. Successivement : normal, coupure à 5 000, 2 000 et 800 Hz. [M. C.]

Son 3.2 – Voix parlée : filtrage passe-haut. Successivement : normal, coupure à 1 000, 2 000 et 6 000 Hz. Fin en voix normale. Texte extrait de *Zadig* de Voltaire. [M. C.]

5.2. Bandes critiques, asymétrie du masquage (écoute au casque conseillée)

Son 3.3 – Comparez l'intensité des bruits de bande. Vous allez entendre huit paires de bruits. Le premier bruit de chaque paire, centré à 1 000 Hz, a une largeur de bande constante de 15 % (930 à 1 075 Hz) ; le second, initialement centré sur la même fréquence, a une largeur de bande qui croît à chaque nouvelle présentation, mais son amplitude diminue de façon que la puissance reste constante. À partir de quelle présentation entendez-vous une variation d'intensité pour le deuxième bruit ? Test réalisé par Houtsma A. J. M., Rossing T. D., Wagenaars W. M., 1987. CD *Auditory demonstrations*, piste 7. [ASA-IPO[39]]

Son 3.4 – Asymétrie du masquage d'un son pur par un autre. Soit A un son d'intensité constante et B un son dont l'intensité initiale est identique à A puis diminue par paliers de 5 dB à chacune des 10 présentations (excepté au tout début où l'intervalle est de 15 dB). Le test comporte 10 séquences. Chacune est constituée de l'alternance : A (200 ms), silence (100 ms), A + B (200 ms), répétée quatre fois. Une séquence dure 2,4 secondes. Premier test, Son 3.4a : A = 1 194 Hz ($ré_5$) ; B = 1 990 Hz (do_5). Au cours du test, le son aigu B faiblit jusqu'à être masqué. Deuxième test, Son 3.4b : A = 1 990 Hz, B = 1 194 Hz. Au cours du

39. Reproduit avec l'aimable autorisation des auteurs et de l'A.S.A. Copyright, Acoustical Society of America.

test, le son B, plus grave, résiste au masquage. Pour chaque test, comptez le nombre de présentations pendant lesquelles vous pouvez réellement entendre le son masqué, avant qu'il ne disparaisse perceptivement (attention à l'autosuggestion !). CD *Auditory demonstrations*, piste 22. [ASA-IPO, *ibid.*]

5.3. Seuils différentiels

Son 3.5 – Seuil différentiel d'intensité (à écouter au casque). Dans cet exemple, vous entendez un son sinusoïdal tenu, de 800 Hz (sol_4) dont la partie centrale peut présenter une variation d'intensité croissante. Écoutez attentivement les cinq premiers sons et notez celui pour lequel vous commencez à percevoir cette variation. [Suva Pro, 1997]

Son 3.6 – Seuil différentiel en fréquence. Testez votre capacité à détecter la plus petite variation de fréquence. Le test est constitué de 10 séquences, chacune comportant 4 paires de sons, donc 4 intervalles. Le premier son de chaque paire est toujours à 1000 Hz ; le second son est soit plus haut, soit plus bas. L'écart, qui est de ± 10 Hz à la première paire, diminue progressivement à chaque nouvelle séquence. Notez vos réponses à l'aide d'un code, par exemple A pour un intervalle ascendant et D pour un intervalle descendant. Pour la première séquence, vous devez avoir comme réponse : A, D, A, A. CD *Auditory demonstrations*, piste 33 [ASA-IPO, *ibid.*]

5.4. Finesse de discrimination auditive

Son 3.7 – Début de *l'Octuor* de Mendelssohn. L'exposition du thème fugué est d'abord entendue au violoncelle seul, deux fois de suite. La discrimination mélodique est difficile. [Archives LAM]

Son 3.8 – Entrée successive des instruments, du grave à l'aigu. Au fur et à mesure que la tessiture monte, le thème devient de plus en plus distinct. [*Id.*]

5.5. Variation de la sensibilité auditive avec la fréquence

Son 3.9 – Trois sons purs enregistrés au même niveau sonore. Successivement : 100 Hz ; 1000 Hz ; 10000 Hz. Remarque : au-delà d'un certain âge, on n'entend plus le dernier son (voir figure 3.17). [M. C.]

5.6. Durée du son et sensation de hauteur tonale

Son 3.10 – Dans cet exemple, vous pourrez apprécier les changements de perception liés à la durée du son. Initialement la durée du son est d'une période. Cette durée double à chaque présentation. La sensation est d'abord celle d'un bruit, d'un clic, et se transforme progressivement en hauteur musicale. Selon les auditeurs, on peut entendre un changement descendant, ascendant, et percevoir la « note » plus ou moins tôt. Comparez vos résultats en fonction de la fréquence. Dans cet exemple, le test porte sur un son de 300 Hz. CD *Auditory demonstrations*, piste 29. [ASA-IPO, *ibid.*]

Son 3.11 – Même expérience avec un son de 1000 Hz. [*Ibid.*]

Son 3.12 – Même expérience avec un son de 3000 Hz. [*Ibid.*]

5.7. Perception mélodique dans l'aigu et plafond du codage temporel

Son 3.13 – Écoute de sons purs dont on double la fréquence. On entend successivement 8 couples de sons : 63-125, 125-250, 250-500, 500-1000, 1000-2000,

2000-4000, 4000-8000, 8000-16000 Hz. Ces octaves, justes du point de vue numérique, peuvent être perçues trop grandes ou trop courtes selon la tessiture et différemment selon les auditeurs. [M. C.]

Son 3.14 – L'exemple sonore comprend 5 mélodies courtes construites sur 5 notes. Il faut écouter en premier la mélodie 3.14a qui se situe dans une tessiture très aiguë et la transcrire graphiquement, soit à l'aide de traits ou de points reliés entre eux, soit sur une portée musicale. *Ensuite seulement* écouter la mélodie 3.14b, un peu plus basse, et la noter sur une nouvelle feuille. La tâche devient de plus en plus aisée au fur et à mesure que la tessiture s'abaisse. Les limites d'ambitus des mélodies sont successivement: 1, (4800-7520) ; 2, (3600-5640) ; 3, (2400-3760) ; 4, (1200-1880) ; 5, (600-940). Selon les données de la psychoacoustique, les mélodies 1 et 2 ne devraient pas être perçues musicalement ; autrement dit, les intervalles ne seraient pas appréciables, or certains auditeurs y parviennent. [M. C.]

5.8. Plusieurs exemples de sons dont les composantes sont équidistantes de 250 Hz

Son 3.15 – Sur une même fondamentale, 250 Hz, on entend ; Son 3.15a, son pur ; Son 3.15b, son harmonique ; Son 3.15c, son harmonique privé des deux premières composantes ; Son 3.15d, son ne comportant que les harmoniques 9-10-11 ; Son 3.15e, son dont les trois composantes, distantes de 250 Hz, sont inharmoniques. Voir la figure 3.28. [M. C.]

5.9. Discrimination des harmoniques

Son 3.16 – Cet exemple fait entendre un son synthétisé avec 5 harmoniques. Partant de do_1 = 65,4 Hz, il est ensuite transposé sur do_2 (130,8 Hz), do_3 (261,6 Hz), do_4 (523,2 Hz), do_5 (1046), do_6 (2092). Les harmoniques sont aisément séparables sur les premières octaves. On peut remarquer qu'à partir du do_4 la discrimination se perd graduellement. [M. C.]

5.10. Du discontinu au continu

Son 3.17 – Suite d'impulsions accélérant progressivement. Au-delà d'une certaine rapidité (seuil temporel), on perçoit une hauteur continue ascendante. [M. C.]

5.11. Modulations et battements

Son 3.18 – Modulation d'amplitude d'un son de 525 Hz (environ do_4). La fréquence de modulation croît progressivement de 1 à 82 Hz en produisant l'effet d'un accord complexe. [M. C.]

Son 3.19 – Modulation de fréquence d'un son de 525 Hz. La fréquence de modulation croît de 1 à 82 Hz en produisant aussi un accord complexe final. [M. C.]

Son 3.20 – Modulation de fréquence : effet de la vitesse de modulation. Sur le même son de 525 Hz, on applique successivement les modulations de 4, 6, 8 et 12 Hz. [M. C.]

Son 3.21 – Sifflet à roulette (sifflet d'agent de police). [LAM]

Son 3.22 – L'effet d'une modulation rapide de 82 Hz produit un accord musical. Comparaison des modulations d'amplitude et de fréquence de même rapidité, 82 Hz, sur une porteuse sinusoïdale de 525 Hz. [M. C.]

Son 3.23 – Comparaison d'une modulation d'amplitude et d'un battement : sons purs. On entend successivement un son pur de 525 Hz modulé à 4 Hz et le

battement de 4 Hz résultant de la combinaison de deux sons purs, 523 et 527 Hz. [M. C.]

Son 3.24 – Même expérience que le Son 3.23 mais réalisée cette fois avec des sons complexes, riches en harmoniques. Modulation et battement sont nettement distincts. [M. C.]

Son 3.25 – Battements lents de fréquence constante affectant trois sons de fréquence différente : 220 Hz, 440 Hz et 880 Hz. La séquence fait entendre deux rapidités de battements : d'abord 4 Hz, puis 8 Hz. La sensation due au battement reste assez stable, tant que l'on peut suivre la modulation d'amplitude ; expérience à comparer avec celle du Son 3.26. [M. C.]

Son 3.26 – Même séquence que précédemment avec des battements rapides : 30 Hz et 60 Hz. Cette fois, la modulation due aux battements change considérablement avec la tessiture. Dans le grave, les battements sont perçus comme un intervalle – qui change selon la tessiture – alors que dans l'aigu ils produisent un effet de rugosité variable. [M. C.]

Son 3.27 – Un paradoxe de l'écoute binaurale : les variations de sonorité à l'enregistrement. Son 3.27a : deux notes jouées à la flûte traversière, enregistrées en stéréo. Son 3.27b : même exemple que précédemment mais mono ; successivement la voie gauche puis la voie droite. L'écoute attentive de la première note permet de remarquer une importante différence de sonorité entre les deux captations. Son 3.27c : montage de la première note isolée, d'abord voie gauche, puis voie droite. Comparez l'écoute en champ libre et l'écoute au casque. [M. C.]

6. Réponses aux tests

Son 3.3, bandes critiques. Dès la deuxième paire, le deuxième bruit a une bande passante plus large que celle qui est estimée pour la bande critique centrée à 1000 Hz. Mais vous pouvez tout aussi bien ne percevoir une différence qu'à partir du 3e ou du 4e couple, d'autant que les changements de spectre dus à l'élargissement de la bande passante induisent des changements de hauteur perçue qui perturbent quelque peu l'estimation de l'intensité !

Son 3.4, asymétrie du masquage. Ce test démontre clairement que les sons aigus sont plus aisément masqués que les sons graves. Attention : ceci n'est valable que pour des sons purs.

Son 3.5, seuil différentiel d'intensité. La variation d'intensité est successivement de 0 ; 0,5 ; 1 ; 2 ; 3 ; 6 ; 10 ; 20 ; 40 ; 60 dB. Il est courant de ne la détecter que pour les numéros 3 ou 4.

Son 3.6, seuil différentiel de fréquence à 1000 Hz. Les dix groupes, [1] A, D, A, A ; [2] A, D, D, D ; [3] D, A, A,D ; [4] D, A, A, D ; [5] A, D, A, D ; [6] A, D, A, A ; [7] D, D, A, A ; [8] A, D, A, D ; [9] D, D, D, A ; [10] D, A, A, D. L'écart diminue de 1 Hz à chaque présentation. Pour le groupe 10, il n'est plus que de 1/1000 de hertz, ce qui correspond à un intervalle de 1,6 cent ou 0,4 savart.

CHAPITRE 4

UNE APPROCHE DE LA PERCEPTION SONORE : FORMES ET CATÉGORISATION

1. La perception

1.1. Préambule

L'enregistrement, et surtout la synthèse sonore, ont opéré un tournant décisif dans notre rapport aux sons en nous donnant la maîtrise de la matière sonore. De même que Lavoisier put faire la preuve que l'eau était constituée d'une certaine combinaison d'atomes d'hydrogène et d'oxygène, la réalisation des premiers générateurs électriques montra que l'on pouvait créer des sons en combinant différemment fréquences, phases et amplitudes des circuits oscillants. Sur cette base de connaissances se sont développées des techniques d'analyse et de contrôle des sons, des instruments de musique de synthèse et, dans le même mouvement, une théorie de la perception auditive consistant à rendre compte de la façon dont un auditeur humain capte et interprète les paramètres physiques du son. Le fait que la musique, domaine d'expertise du son, soit représentable dans notre culture sous la forme graphique d'une partition écrite, donc paramétrable, a longtemps conforté cette vue. Toutefois, pour aborder la perception sonore dans son ensemble, examinons plutôt comment nous réagissons quotidiennement aux sons de notre environnement.

Que le lecteur veuille bien faire maintenant une pause dans sa lecture et porter attention aux différents sons qui lui parviennent en ce moment même. Qu'entend-il ?

Selon qu'il est à la maison (en ville ou à la campagne), dans un train, seul ou en compagnie, il pourra dire qu'il entend « des voitures, le roulement du train, un oiseau qui piaille, des conversations, etc. ». L'emploi de ces expressions pour décrire ce qu'il entend nous indique déjà qu'il est spontanément porté à rendre compte de ce qui se passe, plutôt que de se livrer à une analyse des différents paramètres du son ! Si tout signal sonore est une combinaison de fréquences, d'amplitude et de phases, c'est avant tout pour un auditeur un événement porteur de sens.

Il y aurait donc d'un côté la matière sonore, le *signal* qu'analyse et maîtrise l'acousticien, et d'un autre côté les *sons* que nous percevons comme autant d'entités reconnaissables. Jeter un pont de l'un à l'autre est l'enjeu d'une étude de la perception sonore.

1.2. Qu'est-ce que percevoir ?

La question peut sembler naïve. Percevoir, c'est voir, entendre, toucher, sentir, goûter, c'est user de nos sens pour explorer à la fois notre propre corps et ce qui l'entoure. En poursuivant plus avant la réflexion, on découvre rapidement que l'analyse des relations qui s'établissent entre un sujet percevant et les objets du monde se révèle plus complexe que le simple bon sens ne le laisse entrevoir. En introduction d'un ouvrage intitulé *La perception, essai sur le sensible*, Renaud Barbaras écrit :

> *La perception est (donc) caractérisée par une double dimension. D'un côté, elle est un mode d'accès à la réalité telle qu'elle est en elle-même ; dans la perception, je n'ai à aucun moment le sentiment d'avoir affaire à un double, à une image de la chose : j'ai au contraire la conviction de découvrir une réalité qui précède mon regard et telle qu'elle était avant que je la perçoive. De l'autre, cependant, la perception est sensible, c'est-à-dire mienne : elle est l'épreuve que je fais de la réalité. On traduit ainsi le fait incontestable que, sans sujet percevant, précisément sans organe des sens, rien n'apparaîtrait.*
>
> Barbaras, R., 1994, p. 3.

C'est le sens même de l'acte de perception que la tradition philosophique interroge ici : s'agit-il de la réception passive des données du monde extérieur ou d'un processus actif d'élaboration d'une « représentation » de ce monde à travers nos sensations ? Bien que schématique, la question ainsi formulée représente assez bien les grands courants de pensée qui font l'objet de débats philosophiques – citons Aristote, Locke, Descartes, Kant, Bergson et Husserl pour ne mentionner que les auteurs les plus marquants – et qui sous-tendent aujourd'hui les recherches sur la perception. Celles-ci sont extrêmement diverses et s'appuient encore largement sur des méthodes expérimentales nées à la fin du XIXe siècle, époque pendant laquelle le puissant courant de développement des connaissances scientifiques et artistiques, issu de l'*Encyclopédie*, se tourne vers l'exploration du vivant. Il s'agit de mieux comprendre les relations qui lient la mesure des grandeurs physiques (distance, poids, luminosité, etc.) à nos estimations sensorielles. L'étude de la vision et du toucher prend beaucoup d'avance sur celle de l'audition, totalement tributaire des techniques de production et de contrôle des sons qui ne se développent qu'à la fin du XIXe siècle. L'invention du téléphone (Graham Bell, 1876), en particulier, suscite de nombreuses recherches sur le sujet, car il devient utile de mieux connaître les propriétés statistiques des oreilles humaines pour réaliser une bonne adaptation des récepteurs[1]. De concert avec les avancées des sciences physiques, avec celles de l'anatomie des récepteurs sensoriels (Helmholtz pour l'oreille et pour l'oeil) et de la physiologie du système nerveux, se développe une psychophysique sensorielle dont le traité de Fechner[2] marque la naissance. L'étude des sensations humaines est entrée au laboratoire. En parallèle se développe une psychologie expérimentale qui soumet nos capacités les plus nobles – la mémoire, la pensée, l'intelligence, le langage – à une étude objective et à des évaluations quantitatives. En 1931, Paul Guillaume écrit dans son ouvrage *Psychologie*, destiné aux élèves de la classe philosophie : « On se propose, à l'exemple des sciences de la nature, de décrire des *faits* et de déterminer leurs *conditions*, c'est-à-dire d'autres faits dont l'observation montre le rapport constant avec les premiers. » Plus tard (1960), alors qu'il revendique la psychologie comme recherche indépendante, il

1. L'adaptation séculaire des sources sonores traditionnelles à l'oreille humaine s'est toujours réalisée par une succession d'essais et d'erreurs, en particulier dans le domaine musical, mais cette fois l'enjeu commercial requiert des recherches systématiques en laboratoire.
2. Gustav Fechner, 1860, *Elemente der Psychophysik*, Leipzig : Breitkopf & Härtel. Voir aussi le chapitre 3, § 2.4.

ajoute : « en d'autres termes, on se propose d'établir des *lois* »[3] (*Manuel de psychologie,* p. 3).

D'autres auteurs soulignent que, par comparaison aux instruments de mesure de la physique, nos organes des sens paraissent bien imparfaits : imprécis, peu stables, sujets à illusion et fournissant des données variables dans le temps.

> *Nous pouvons maintenant nous rendre compte que nos appareils sensoriels ne constituent pas des fenêtres s'ouvrant sur le monde extérieur, et qu'enfermés dans notre caverne subjective nous n'observons même pas les ombres des passants qu'invoquait le symbole platonicien.*
>
> Piéron, H., 1945, p. 412.

Ce courant de recherche s'inscrit dans une démarche objective, rigoureuse, qui va du simple au complexe, partant des sensations élémentaires pour aboutir aux perceptions élaborées. Censé rendre compte de la réception sensorielle dans sa généralité, il opère principalement avec des stimuli élaborés en laboratoire afin que les paramètres physiques en soient contrôlables à tout moment. Le souci d'objectivité qui oriente une recherche expérimentale quantitative et le fait que les résultats s'expriment sous forme numérique confèrent en quelque sorte une légitimité indiscutable au courant psychophysique.

Pourtant, d'autres chercheurs, non moins soucieux de rigueur et d'objectivité, se posent la question du bien-fondé de l'emploi de telles méthodes pour évaluer les capacités perceptives des êtres vivants. Et si les caractéristiques nécessaires aux appareils de mesure (précision, fiabilité, reproductibilité) se révélaient inutiles, voire invalidantes, pour explorer un environnement variable tel qu'il se présente dans la vie ?

Dès la fin du XIXe siècle naissent des courants dissidents. En réaction aux « élémentaristes » qui cherchent à décomposer les perceptions en sensations unitaires, se développe la théorie de la forme (Von Ehrenfels, 1890 ; Wertheimer, 1925) qui rejette la distinction classique entre sensation et perception[4].

Plus récemment, James Gibson, psychologue *écologique* (1979), affirme que la fonction perceptive des êtres vivants ne peut être étudiée indépendamment du lieu naturel d'activité. C'est par la perception qu'un être vivant peut se positionner dans le monde, échapper aux dangers, trouver sa nourriture, ce qui nécessite de réagir rapidement à des stimulations reconnaissables sans ambiguïté et de sélectionner celles qui sont pertinentes pour lui.

Étant donné que les situations sont changeantes, que les événements sonores ne se reproduisent jamais deux fois de la même façon, il apparaît que l'aptitude à reconnaître des formes similaires l'emporte sur la mesure des paramètres du signal, car il est plus efficace d'estimer globalement la similarité entre deux phénomènes que d'effectuer l'analyse exhaustive des propriétés de chacun d'entre eux.

Aujourd'hui, les principales caractéristiques des deux courants de recherches peuvent se résumer ainsi.

3. Nous avons conservé les termes en italique dans le texte de Guillaume.
4. Voir André Delorme, *Psychologie de la perception*, 1982, p. 30.

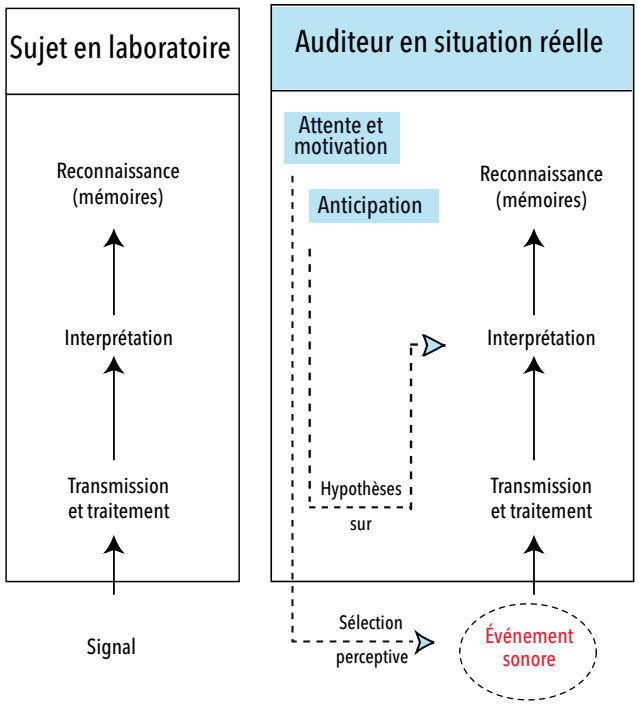


Figure 4.1 Deux conceptions de la perception. À gauche, schéma *bottom-up* ou ascendant. À droite, schéma *top-down*. La partie descendante (flèches bleues) précède ou court-circuite la partie ascendante.

Pour les uns, que nous qualifierons de *physicalistes*, le monde extérieur constitue une réalité exacte dont chaque individu ne capte qu'une faible partie par le moyen d'organes des sens aux capacités limitées et aux performances moins stables et moins précises que les capteurs physiques dont nous disposons (microphone, caméra). La perception se construit essentiellement dans un sens ascendant (figure 4.1, gauche), ce qui implique une progression partant des sensations les plus simples – réponses sensorielles aux paramètres du stimulus – vers des traitements plus élaborés, pour aboutir finalement à l'interprétation par le cerveau.

Pour les autres, que nous nommerons *cognitivistes*, percevoir, c'est construire une représentation du monde utile pour la survie, propre à chaque individu, qui s'enrichit, se complexifie, se stabilise au cours des interactions multisensorielles avec l'environnement extérieur, et dans l'échange avec les autres membres de la société. L'approche cognitiviste privilégie la projection descendante (figure 4.1, droite) de données mémorisées dans le cerveau vers les organes des sens, à seule fin de ne saisir, dans la complexité du monde, que les éléments pertinents pour l'individu. Ce schéma implique que la perception est sélective et, pour agir efficacement sur le monde extérieur, qu'elle comporte une importante part d'anticipation.

Les deux démarches sont complémentaires. Il est clair que l'approche dite physicaliste ou *bottom-up* a bien pour objectif d'expliquer la reconnaissance des formes et devra, à terme, prendre en compte les données culturelles du sujet ainsi que l'anticipation constitutive de tout acte vivant. Il est clair aussi que l'approche qui privilégie les traitements cognitifs, ou *top-down*, doit s'inscrire dans la physiologie, depuis l'organe auditif jusqu'aux aires corticales. Actuellement, les deux démarches

font appel à des méthodes de recherche distinctes et sont développées dans des communautés de chercheurs différentes.

Une terminologie à préciser. Il faut mettre ici en garde le lecteur sur le fait que des termes comme *perception, information, forme*, revêtent des sens et des contenus souvent fort différents d'un ouvrage à l'autre. Pour la plupart des chercheurs, et en particulier pour les physicalistes s'inscrivant dans une tradition totalement objectivable et mécaniste, la perception sonore recouvre l'ensemble des opérations biomécaniques et neurophysiologiques déclenchées par une stimulation sonore, depuis l'ébranlement du tympan jusqu'au cerveau. Les étapes de la perception sont observables expérimentalement, en laboratoire. La signification que le sujet attribue aux signaux sonores, qui en est l'étape ultime, n'est généralement pas prise en compte. Le chapitre 3 rend compte des connaissances produites par les recherches psychophysiques sur le système auditif.

Dans la suite de ce chapitre, nous développons l'approche dite écologique qui part du sujet, de l'acte d'un vivant :

> *Le point de départ n'est donc pas la réalité absolue, postulat naïf du sens commun, mais le sujet connaissant, ou plutôt l'agent de l'acte de penser, de sentir, de percevoir, de croire, qui est aussi celui de l'acte de se décider et d'agir.*
>
> Berthoz, A., & Petit, J.-L., 2006, p. 93.

Bien que cette approche prenne en compte l'ensemble des modalités sensorielles, c'est la perception visuelle qui est de loin la plus étudiée, et qui sert généralement de modèle pour les autres. Or, la perception sonore, de même que celle du mouvement, a pour particularité de traiter d'« objets temporels », difficiles à saisir et à formaliser, dont les caractéristiques intrinsèques conditionnent l'ensemble de l'organisation perceptive.

Le terme « perception » sera pris ici dans un sens général incluant l'ensemble des processus qui concourent à la manifestation consciente, pour un individu, de l'existence d'un monde extérieur dont il (se) constitue une « représentation » qui lui est propre.

2. À l'écoute du monde sonore environnant

2.1. L'audition « sens d'alerte » et l'écoute choisie

Nous évoluons dans un univers traversé de conversations, de bruits de circulation, de musique, de signaux divers qui nous sont tellement familiers que nous les entendons sans y prêter attention. Il suffit qu'un son inhabituel ou incongru se produise pour qu'immédiatement nous soyons sur le qui-vive : preuve que l'audition, sens que nous ne pouvons pas « déconnecter », est sollicitée sans relâche. La fonction première de l'audition est d'être un sens d'alerte efficace qui nous informe à tout moment sur les événements pouvant se produire en dehors de la vue, pendant les moments d'inattention, pendant le sommeil. Les deux interrogations qui surgissent alors sont « Où ? » et « Quoi ? », puisque comprendre ce qui se passe nécessite de pouvoir déterminer rapidement la provenance et la nature d'un danger éventuel. La réponse à la première question est réflexe : nous tournons immédiatement la tête dans la direction du son, ce qui a pour effet d'équilibrer l'intensité sonore sur les deux oreilles et d'orienter la vue vers la source. La réponse à la deuxième question est l'objectif immédiat de la perception : identifier la cause du son, lui attribuer un « sens ».

Or, un son quelconque est toujours signe qu'il se passe quelque chose, qu'une action s'est produite ou est en train de se produire. Il peut être dû à un déséquilibre survenant dans le monde physique – intempérie, chute d'objet – ou à une manifestation de la vie, car la vie est mouvement et tout mouvement produit du son[5]. Action et mouvement inscrivent d'emblée la perception sonore dans la temporalité. Ce point fondamental distingue la modalité sonore de la modalité visuelle généralement prise comme référence dans l'étude de la perception. Nous pouvons revenir sur nos pas pour regarder de plus près une affiche, ce qui n'est jamais possible avec le son. Dès qu'il cesse, le phénomène qui a déclenché notre perception n'existe plus que dans notre souvenir.

Pour la plupart d'entre nous, les sons produits par la circulation urbaine, qui sont d'une grande prévisibilité, ne retiennent pas notre attention : ils constituent le « bruit de fond ». Pourtant John Cage déclare :

> *Ma musique : les sons d'ambiance de l'environnement. J'habite la Sixième Avenue ; la circulation y bat son plein. Résultat : à tout instant, une profusion sonore.*
>
> Cage, J., 1994, p. 102.

De façon provocatrice, le compositeur souligne ici le fait qu'à tout moment nous pouvons aussi choisir d'écouter de tels sons, non plus pour leur identité mais pour leurs qualités. Ainsi en est-il d'un chant d'oiseau, d'une voix, de l'ambiance sonore du marché qui nous invitent au plaisir de voyager dans les sons, d'y découvrir des sensations imprévues, sans limitation de temps, sans la contrainte d'une adéquation logique avec la situation.

À l'écoute événementielle du « Quoi ? », requise par la recherche de l'identité des sources, s'oppose donc celle du « Comment ? », c'est-à-dire l'écoute hédonique des qualités librement décidée par l'auditeur. Paul Guillaume voit des orientations contraires dans ces deux modalités perceptives :

> *Nous avons vu surtout, dans l'éducation des sens, un progrès dans la perception des objets et de leur signification. Mais parfois elle suit une direction contraire. Orientée vers des fins esthétiques, elle tend à faire disparaître la « signification » au profit de la « qualité sensible ».*
>
> Guillaume, P., 1931, p. 183.

2.2. Deux modalités perceptives distinctes : identifier les sons ou les qualifier

Quelles sont dès lors, les relations entre l'écoute causale, celle qui consiste à reconnaître le mode de production des sons, et l'écoute qualitative que nous venons de mentionner ? Sommes-nous pleinement libres de décider de l'une ou de l'autre ? En guise de réponse, nous proposons au lecteur de se prêter à l'écoute du Son **4.1**, muni d'un crayon et d'un papier pour répondre à la consigne suivante :

Son 4.1 (28")

Consigne d'écoute

Décrivez les qualités sonores de l'instrument.

Dès le début de l'audition, un trouble, une gêne se manifestent, car il est vraiment difficile de rapprocher le son entendu de celui d'un instrument connu. On constate que l'attention est mobilisée par les tentatives d'identification de la source instrumentale, au détriment d'une analyse qualitative proprement dite. D'ailleurs, à l'exception de quelques-uns (notamment les compositeurs et les auditeurs de musiques

5. Seuls les anges passent en silence !

électroniques ou acousmatiques qui ont l'expérience des transformations du son), les auditeurs répondent rarement à la question posée. Selon leur expertise et leur bagage culturel, ils se prêtent à des commentaires différents, mais, pour tous, l'attribution d'une identité instrumentale est la première réaction d'écoute (pour plus de détails, voir chapitre 7, § 4.1).

Il semble donc qu'il n'y ait pas réellement d'indépendance entre les deux modalités de perception sonore que sont l'écoute causale « Quoi ? » et l'écoute qualitative « Comment ? » La recherche d'une causalité paraît prioritaire bien que nous n'en ayons habituellement pas conscience car les sons qui nous entourent sont en très grande partie prévisibles. Il faut qu'une ambiguïté ou une incongruité se produise pour que nous soyons subitement interpellés. Ainsi, le seul fait de s'interroger sur le genre d'une voix entendue à la radio (un homme ou une femme ?), peut faire perdre momentanément le sens des propos. Comme nous venons de le voir, l'écoute musicale n'y échappe pas. L'introduction d'un instrument inconnu perturbe les auditeurs portés à apprécier les qualités sonores d'une interprétation musicale ou de la structure d'une œuvre. L'antagonisme entre ces deux modes de traitement des sons peut aussi expliquer les difficultés de réception de musiques réalisées avec des sons issus de l'environnement plutôt qu'avec des instruments de musique. Nous y revenons abondamment dans le chapitre 7, à propos de l'étude du timbre (voir chapitre 7, § 1.7).

L'identification[6] et la qualification du son diffèrent aussi dans la durée du traitement perceptif. Il suffit d'écouter le Son **4.2** pour s'en convaincre : en une fraction de seconde, nous avons reconnu l'instrument et nous devenons aptes à porter un avis sur ses qualités.

Son 4.2 (38'')
Original (début)

À l'opposé, apprécier les qualités du son d'un instrument ou d'un enregistrement requiert une durée d'écoute suffisante, car c'est une opération complexe qui exige une certaine expérience pour analyser les paramètres du son : justesse, dynamique, qualités spectrales. Il faut que l'exemple sonore nous laisse le temps d'effectuer différentes comparaisons mentales : au besoin, nous demandons une deuxième écoute.

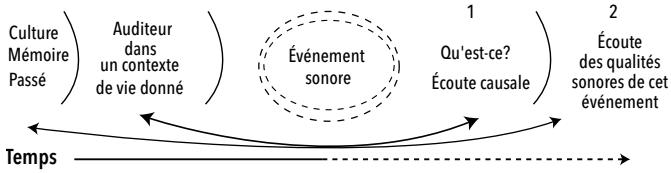


Figure 4.2 La reconnaissance précède la qualification des sons.

Cette expérience met aussi en lumière le fait que nos capacités à porter des jugements sur les qualités se portent habituellement sur les sons de sources connues. L'immense majorité des expériences de perception sonore a, jusqu'à une époque récente, utilisé exclusivement des sons périodiques de synthèse dits musicaux, se prêtant aisément à la paramétrisation des dimensions acoustiques (fréquence, intensité, spectre), avec une correspondance implicite aux dimensions musicales de hauteur, de nuance et de timbre. Cependant, nous verrons que pour être valablement corrélée à la perception musicale l'interprétation perceptive des paramètres spectraux du signal acoustique doit toujours être rapportée à une source identifiée, une production musicale sur un instrument – réel ou de synthèse – connu de l'auditeur.

6. On parle aussi de la « reconnaissance » : dans cette présentation de la perception sonore, nécessairement réduite, nous ne nous attarderons pas sur la distinction entre ces deux concepts, étant entendu qu'il s'agit dans les deux cas d'un processus global impliquant les connaissances du sujet.

Nous poserons donc que l'écoute causale, qui est vitale, précède temporellement l'écoute qualitative, laquelle n'opère valablement que sur des sons reconnus (figure 4.2). Sur ce point, notre approche de l'étude perceptive des qualités des sons diffère de l'approche psychophysique qui stipule que la reconnaissance des sources est l'aboutissement d'un processus d'analyse et d'interprétation des attributs de qualité des sons.

2.3. L'acte d'écoute : sensation, mémoire, anticipation

2.3.1. Analyse de l'écoute d'une mélodie par Husserl

> *La chose semble tout d'abord fort simple ; nous entendons la mélodie, c'est-à-dire nous la percevons, car entendre, c'est percevoir. Pendant que résonne le premier son, le second arrive, puis le troisième, etc. Ne devons-nous pas dire : quand le second son résonne, alors je l'entends lui, mais je n'entends plus le premier, etc. ? En vérité je n'entends donc pas la mélodie, mais seulement le son individuel présent.*

Husserl termine son analyse par une phrase remarquable :

> *Je n'entends donc à chaque fois que la phase actuelle du son, et l'objectivité de l'ensemble du son qui dure se constitue dans le continuum d'un acte qui, pour une part, est souvenir, pour une part, très petite, ponctuelle, perception*[7]*, et pour une part plus large, attente.*
>
> Husserl, E., 2002, p. 36, [le texte complet est reproduit à l'annexe D].

À l'évidence, l'écoute d'une simple mélodie engage le sujet dans un processus actif de gestion en temps réel de la sensation sonore, impliquant deux fonctions essentielles : la mémorisation et l'anticipation. Il en est ainsi pour toute séquence sonore, avec seulement des différences d'échelles dans le traitement temporel et des différences de complexité dans les sollicitations de la mémoire.

2.3.2. Les mémoires

Le mot « mémoire » évoque d'emblée la capacité que nous avons à retenir les connaissances apprises, les émotions, les événements qui se sont produits dans notre vie passée et sur la base desquels s'est construite notre identité. Il s'agit de la mémoire permanente, dite à long terme, dans laquelle sont stockées les règles de syntaxe de la langue et celles de l'organisation des sons dans une musique donnée. Pour comprendre la perception, il faut aussi faire appel à des capacités de mémorisation temporaires, dites à court terme, qui permettent, comme l'évoque Husserl, de garder le souvenir du son que nous sommes en train de percevoir pendant la perception du son suivant.

La mémoire à court terme est la capacité qui nous permet de retenir pendant quelques secondes un numéro de téléphone ou le début d'une phrase parlée. Elle est d'une importance capitale pour l'écoute et la compréhension de la musique. Divers exercices de « dictées musicales » ont d'ailleurs pour objectif d'en développer l'empan et la précision. Les chercheurs distinguent aujourd'hui deux étapes dans le processus de mémorisation à court terme[8] : la première est sensorielle et la seconde, syntagmatique. Un peu d'attention à l'écoute d'un son nous fait prendre conscience qu'une trace vivante quasi exhaustive de l'impression produite persiste pendant quelques fractions de seconde, puis s'évanouit. Il ne subsiste ensuite que ce qui a pu être interprété, codé comme mot ou comme accord.

7. Nous avons déjà signalé que le mot « perception » revêt des sens différents selon les auteurs. Dans cette traduction d'Husserl, « perception » désigne l'activité physiologique de la réception sensorielle.
8. Certains auteurs comme Lindsay et Norman considèrent qu'il s'agit de deux encodages différents et emploient deux termes distincts : *registre d'intégration sensorielle* pour le premier, *mémoire à court terme* pour le second. Pour la plupart des auteurs, la mémoire à court terme désigne l'ensemble des processus qui se déroulent pendant la période de rétention, soit environ les quinze premières secondes. D'autres termes sont en usage comme mémoire *phosphorescente* (Moles), *instantanée* (Leipp), ou encore *échoïque*, pour le son, par comparaison avec *iconique*, pour la vision.

Il faut donc admettre que pendant la durée de rétention se produisent des communications entre « mémoire à court terme » et « mémoire à long terme » comme indiqué sur la figure 4.3.

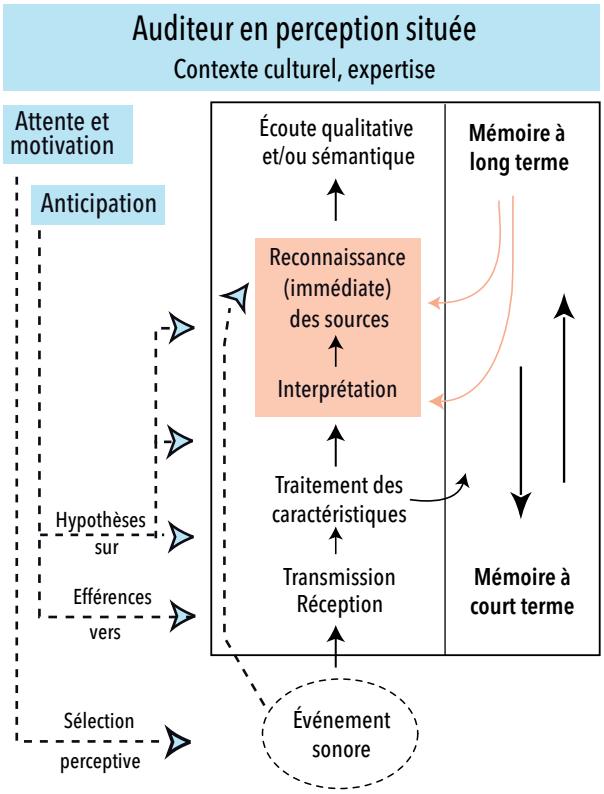


Figure 4.3 Schéma proposant une interprétation des processus de traitement et de mémorisation au cours des étapes de la perception sonore.

À la suite des travaux de Baddeley (1993), les chercheurs considèrent aujourd'hui la mémoire à court terme comme une *mémoire de travail* apte à traiter des tâches différentes et des données pouvant provenir d'autres entrées sensorielles (voir Lecocq, 1993 ; Nicolas, 2003, chapitre 6). Du point de vue temporel, si aucune perturbation ne trouble la rétention, il est possible de conserver ainsi 10 à 15 secondes d'événements sonores, la phase de trace sensorielle proprement dite ne durant que 2 à 3 dixièmes de secondes. Les données numériques que nous avançons ne sont qu'indicatives, puisque la capacité de rétention temporaire dépend de la complexité des événements sensoriels et surtout de l'expérience acquise dans la reconnaissance et l'encodage symbolique (sémiotique) des éléments perçus. Les musiciens, en particulier, développent au cours de leur apprentissage des stratégies d'écoute adaptées aux types de musique et aux styles des compositeurs qui leur permettent de saisir en très peu de temps une grande quantité d'éléments de structure. La composante descendante anticipatrice (en bleu sur la figure 4.3) y est très importante et dirige fortement l'attention sélective grâce à laquelle l'écoute est à la fois riche et efficace. Il y a lieu de remarquer que les éléments de structure que nous venons de mentionner sont de natures très diverses : tour à tour syntaxique (comme

dans l'analyse musicale), qualitative et émotionnelle (comme dans l'appréciation de l'interprétation d'une œuvre connue), acoustique (comme lors de l'écoute professionnelle d'un preneur de son).

2.3.3. L'anticipation : contexte, familiarité, prévisibilité

> *On perçoit le présent à travers les résidus d'expériences « privilégiées » du passé. (...) Toute perception d'objet est une* **sollicitation virtuelle d'actes familiers** *tout prêts à se déclencher. Reconnaître un objet, a-t-on dit, c'est savoir s'en servir. (...) Il faut que la perception, pour être efficace, soit englobée dans une conduite significative, orientée. Dans la plupart des cas, la fonction de la perception est la* **reconnaissance** *des objets d'après quelques signes sommaires, l'***anticipation** *sur l'expérience qui compléterait la connaissance de l'objet. Un grand intérêt vital ou pratique s'attache à cette reconnaissance anticipée qui permet de « réagir à temps », de « préparer » son attitude, « d'économiser » des expériences inutiles ou d'en « éviter » de nuisibles.*
>
> Guillaume, P., 1931, Extraits des pages 177-178[9].

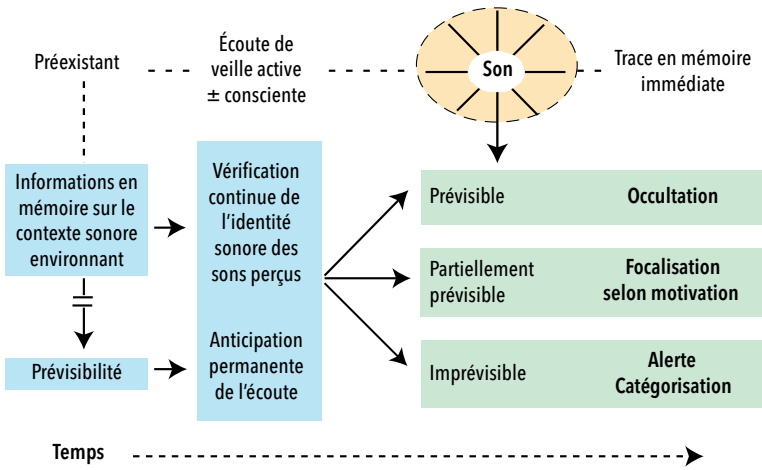


Figure 4.4 Écoute et prévisibilité des sons : l'anticipation.

Remarquons tout d'abord que nous sommes toujours dans un environnement connu : dans la rue, à la maison, dans une salle de cours, au concert, à la plage. Autant de contextes de vie pour lesquels nous avons mémorisé, au cours de nos expériences antérieures, la collection des sons liés aux événements ou aux activités habituelles que la vue nous indique le plus souvent. La plupart des sons que nous entendons sont prévisibles, ils sollicitent peu notre attention et nous pouvons même les occulter totalement. L'écoute que pratique en tâche de fond notre système de vigilance consiste seulement à vérifier de loin en loin que le son perçu correspond bien à l'un des événements sonores susceptibles de se produire dans le lieu où nous sommes (figure 4.4).

La prévisibilité d'occurrence des sons, ou fréquence écologique de Ballas (1999), la connaissance acquise de leur structure acoustique et, à un autre niveau, celle de la syntaxe de séquences sonores comme la parole ou la musique, permettent la construction anticipée du sens qui dispense d'effectuer l'analyse exhaustive du son lorsqu'il se produit. La perception sonore la plus commune est anticipatrice et, par ce fait, économe et efficace.

9. À noter que l'auteur a pris soin de mettre en valeur certains termes : les mots en caractères gras du texte original sont reproduits tels quels et les mots en italiques sont ici entre guillemets.

> *Le cerveau est un prédicteur. S'annoncerait là une révision déchirante du paradigme classique, aux termes de laquelle l'anticipation, au lieu d'une exception, se découvrirait être la vraie règle d'un comportement intelligent.*
>
> Berthoz, A. & Petit, J.-L. 2006, p. 35.

L'anticipation est consubstantielle à la perception sonore. Comment comprendre autrement que nous soyons capables de « traiter en temps réel » la profusion des sons qui nous entourent, soit dans leur identification, soit dans leur qualification ?

L'anticipation permet aussi de configurer à l'avance son écoute pour choisir de concentrer son attention sur des caractéristiques particulières du son comme le font en permanence les musiciens, ou sur des indices décisifs pour la discrimination des bruits comme le pratiquent les « oreilles d'or[10] » de la marine. Il faut effectivement *entendre le son avant* qu'il ne soit effectivement produit pour capter au vol les caractéristiques discriminantes au regard de notre motivation, celles qui constitueront notre interprétation personnelle de l'écoute d'un discours, d'une musique, de l'ambiance sonore d'un environnement.

2.4. L'étude « écologique » de la perception sonore et le paradoxe de la complexité

En prenant pour point de départ un sujet vivant plongé dans un monde réel, la perception apparaît comme une activité fondamentale, celle qui, se fondant à la fois sur les données sensorielles immédiates et sur les connaissances antérieures, donne sens à la situation vécue tout en produisant une bonne anticipation des événements à venir. Pour rester au plus près de l'expérience courante d'écoute, nous prenons le pari, dans cet ouvrage, de travailler avec des sons réels, nonobstant l'obstacle apparent de leur complexité. En effet, les sons de notre environnement quotidien, ceux de la voix humaine, ceux des instruments de musique, sont tous extrêmement complexes au regard de l'acoustique : ils comportent des bruits, des composantes harmoniques ou non et ils sont continûment variables. L'expérience montre qu'ils sont aisés à mémoriser, et ceci en dépit des changements qui les affectent d'une production à l'autre.

> *The number and variety of temporally complex environmental sounds that are identified correctly even by small, inattentive and relatively unskilled children suggests that what is simplest from the sensory researcher's point of view may not be so for the perceiver.*
>
> Vanderveer, N. J., 1979, p. 232[11].

Nancy Vanderveer souligne ici un malentendu fondamental sur la notion de complexité en perception. Partant du principe qu'il est prudent de commencer une étude avec des éléments simples que l'on maîtrise avant d'aborder des phénomènes plus complexes, les acousticiens et les chercheurs sur la perception ont produit un grand nombre de travaux utilisant des sons simples – purs et stables – aisés à produire et à contrôler en laboratoire, et censés, par conséquent, être plus « faciles » à écouter, ce qui est tout à fait contestable. Les sons « simples » permettent d'évaluer l'incidence des différents paramètres d'une stimulation sensorielle. Ils sont l'outil de choix pour étudier les traitements neurophysiologiques des fonctions auditives mais sont étrangers au corpus sonore ordinaire sur lequel nous fondons notre expérience de l'écoute. En prenant le parti de travailler avec des sons calibrés

10. Nom donné aux personnes qui discriminent les bruits des sous-marins dans le bruit de fond de la mer saturé par les crevettes claqueuses, les poissons cloches et les « balanes » bavardes !
11. Le fait qu'un grand nombre et une grande variété de sons de l'environnement, d'une grande complexité temporelle sont parfaitement identifiés même par de jeunes enfants inattentifs et relativement peu expérimentés, laisse entendre que ce qui est le plus simple à expérimenter du point de vue des chercheurs peut paraître difficile à percevoir par les auditeurs. Traduction : M. C.

et reproductibles pour évacuer la complexité apparente, le chercheur court le risque de rester à la marge d'une des caractéristiques la plus remarquable de la cognition des êtres vivants, et plus particulièrement des humains : le traitement de la variabilité. Comme les irrégularités de l'écriture manuscrite, comme les incessants mouvements d'un visage, les sons qui nous entourent ne se produisent jamais deux fois de la même façon, mais la plupart possèdent une stabilité structurelle qu'il s'agit de débusquer.

Par ailleurs, les ondes sonores qui parviennent à nos oreilles portent le mélange de plusieurs sources. Or, nous pouvons très bien discriminer la parole de notre interlocuteur, même si elle est mêlée aux bruits ambiants de la circulation, à d'autres conversations ou encore à une musique de fond. La variabilité et le mélange des sources sonores sont le lot quotidien de la perception sonore.

Une scène sonore

L'exemple sonore **4.3** propose l'écoute d'une scène des plus banales, réduite à la seule dimension acoustique comme le sont les dramatiques radiophoniques. Pour l'avoir expérimentée avec de nombreux étudiants, nous pouvons dire que la scène est intelligible puisque tous sont capables de nommer, dès la première audition, la quasi-totalité des événements sonores qui s'y produisent[12], en les intégrant dans une interprétation globale cohérente[13].

Son 4.3 (17'')

Scène sonore de la vie ordinaire

La reconnaissance des sources et des événements sonores est une tâche tellement banale que nous ne soupçonnons pas la complexité des mécanismes perceptifs qu'elle met en jeu. Ce n'est que lorsque les chercheurs ont été confrontés à l'interprétation des données issues de l'analyse physique, et plus directement lorsqu'on a voulu remplacer l'humain par des systèmes automatiques, que ce problème majeur de la perception qui est la reconnaissance d'objets et leur attribution d'un *sens*, s'est révélé l'obstacle principal à la simulation de la perception humaine. Face aux traitements humains dont l'efficacité et la flexibilité sont étonnantes, les traitements automatiques, bien que puissants, paraissent encore lourds et sujets à des erreurs parfois cocasses, comme l'exprime avec humour Daniel Andler (2002, p. 1087) : « Les jeunes enfants, et les pigeons, sont encore infiniment supérieurs aux machines dans la plupart des tâches de reconnaissance. »

Après les déboires des premières tentatives de segmentation automatique de parole enregistrée, développées dans les années 1970, il est vite apparu que l'identification des unités phonétiques, évidentes pour un auditeur, se faisait d'autant plus aisément que la phrase à reconnaître était déjà en partie connue ! De nouvelles recherches sur les comportements humains de saisie et de traitement de l'information, et en particulier sur les couplages entre les mécanismes de la perception et les connaissances préalables, se sont développées et ont donné naissance aux sciences cognitives. Aujourd'hui, c'est la reconnaissance des sons musicaux qui est devenue un enjeu commercial. Il s'agit d'effectuer l'indexation automatique des enregistrements, ce qui requiert la reconnaissance du timbre instrumental – dont nous verrons qu'il se présente sous un double aspect selon qu'il renvoie à l'un ou l'autre des deux modes d'écoute mentionnés plus haut (voir § 2.2) –, la reconnaissance du style de musique et l'identification de la pièce, voire celle de l'interprète.

Dès lors, comment accéder au support acoustique de cette reconnaissance ? Quels sont les éléments du signal acoustique qui permettent aux auditeurs de discriminer les voix humaines des sons de la musique, d'identifier du liquide, des paroles, une voix de femme et le bruit d'une porte qu'ils n'ont jamais entendus ? Il faut postuler

12. Le bruit d'ouverture d'une porte et son grincement, une voix (toux, parole), des rires (personnes masculines), une musique, une voix féminine disant « merci, etc. », le son d'un bouchon qui saute et les bruits de versement d'un liquide, la voix d'un homme disant « ah, c'était bon », etc.
13. Quelqu'un entre dans une pièce où se trouvent des personnes qui rient, écoutent de la musique et boivent.

qu'un minimum de régularités relient un type de production donné à sa « forme acoustique » perçue, et que les auditeurs apprennent à les repérer et à les mémoriser au travers de la variabilité du monde.

Deux théories développent des concepts permettant d'expliquer comment un sujet vivant plongé dans un environnement sonore complexe, variable, peut repérer des régularités dans le flux sonore continuellement changeant et extraire des éléments stables, mémorisables, qui font sens pour lui, afin d'anticiper ses actes de façon efficace. Ce sont la *théorie de la forme* et la *catégorisation prototypique*.

Les détours que nous allons faire en empruntant des exemples aux sons de l'environnement et à la parole ne nous écartent pas de l'écoute musicale, bien au contraire. Le socle commun à ces trois domaines d'écoute est la reconnaissance de formes acoustiques.

Une forme est un tout

Le principe premier, fondamental, de la théorie de la forme est celui d'unicité de la forme, que l'on exprime en disant qu'une forme est autre chose que la somme des éléments dont elle est constituée : c'est un tout. La grande majorité des exemples illustrant la théorie de la forme proviennent du domaine visuel. Nous empruntons à E. Leipp (1977, p. 20) le lapin de la figure ci-dessous. Sa forme est intégralement réalisée avec les éléments de B mais il faut quelques secondes pour s'assurer que tous les éléments de B sont bien dans A. En combinant différemment ces mêmes éléments on peut réaliser des figures nouvelles. Si quelques lecteurs « voient » instantanément en C un Mexicain à bicyclette vu du dessus, le plus grand nombre reste perplexe à la vue de la figure D. Est-ce une forme ou un agrégat de formes ? Risquons une interprétation : la figure D représente un acousticien asiatique doté d'une oreille gauche hypertrophiée. Aussitôt le déclic s'opère. À présent, chaque forme, bien que constituée des mêmes cinq éléments, apparaît comme une entité distincte : les éléments isolés initiaux ont perdu leur autonomie, nous ne les voyons pas indépendamment de cette forme globale. Un tel jeu n'a pas de limite : nous laissons au lecteur le plaisir de trouver des interprétations pour les formes E et F et de créer d'autres formes.

Cette première expérience révèle immédiatement une difficulté fondamentale de l'étude des formes : est-il possible de caractériser une forme pour elle-même, c'est-à-dire indépendamment du sens que nous pouvons lui attribuer en l'observant ? « Nous ne pouvons jamais distinguer nettement ce que nous voyons de ce que nous pouvons savoir », écrit Gombrich (2002, p. 331). De leur côté, Lindsay et Norman, dont l'ouvrage s'ouvre sur la reconnaissance des formes, posent d'emblée que la perception des formes consiste dans une interaction constante entre le « traitement dirigé par données » (ou *bottom-up*) et le traitement « dirigé par concepts » (ou *top-down*). (Voir page 142)

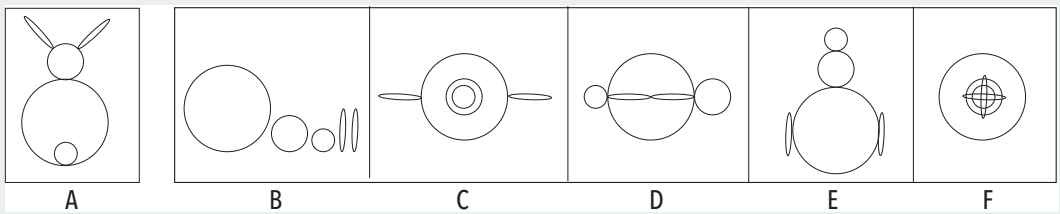


Figure 4.5 Dès qu'une forme prend sens, elle est perçue comme un tout. Les cinq formes élémentaires de B ne sont pas immédiatement perceptibles dans la forme globale de A, qui évoque un lapin, bien qu'ils en constituent la silhouette. En C, on peut reconnaître un schéma classique : celui d'un cycliste à chapeau mexicain, vu du dessus. Le lecteur pourra imaginer des interprétations pour les figures D, E, F.

3. La notion de forme sonore spectrotemporelle : sources et séquences

3.1. Qu'est-ce qu'une forme ?

La notion de forme traverse plusieurs champs disciplinaires parmi lesquels on trouve la philosophie, les sciences de la nature, les sciences mathématiques et l'art. Selon les contextes, la forme s'oppose à la matière, au fond, au sens ou tout simplement à l'amorphe. Le lecteur trouvera un exposé riche et synthétique de la notion de forme dans l'ouvrage déjà cité de Daniel Andler (2002) ainsi que de nombreux exemples illustrés dans les livres de psychologie (Lindsay & Norman, 1980 ; Delorme, 1982), de biologie végétale et animale (D'Arcy Thompson, 1994 ; Stevens, 1978) ainsi que dans des livres d'art (Gombrich, 2002, Baltrusaitis, 1969, Huyghe, 1971).

Ce qui nous concerne ici est la morphologie du sonore dont la dimension structurante est la temporalité et, singulièrement, la question des rapports entre la caractérisation des formes acoustiques et leur perception par les humains.

En 1890, Von Ehrenfels[14] remarque qu'une mélodie n'est pas qu'une simple succession de notes puisqu'elle reste inchangée lorsque toutes les notes ont été modifiées par transposition. La mélodie est donc un tout*, autre chose ou quelque chose de plus que la somme de ses parties*[15] : c'est une forme. La *Gestalttheorie,* ou théorie de la forme, se développe rapidement et s'impose en France grâce à l'ouvrage de Paul Guillaume (1937).

S'il est possible de bâtir une théorie des formes du monde physique, telle la morphodynamique de Jean Petitot[16], il semble bien que l'appréhension des formes par les humains ne soit pas indépendante du (des) sens que ceux-ci peuvent lui attribuer, puisque c'est la finalité même de la perception. Lorsque le Dr Sacks[17] tend une rose à un malade atteint d'une lésion cérébrale gauche, en le priant de dire ce que c'est, ce dernier décrit un objet d'« environ 15 cm de long, une forme rouge enroulée avec une attache linéaire verte ». Le malade ne voit pas de rose. Il discrimine différentes parties et en donne une description topologiquement correcte, mais il a perdu la faculté visuelle de reconnaître la forme globale d'une fleur. Hors de la catégorie fleur, une tige est une « attache linéaire verte ».

La perception d'une forme est une construction cognitive complexe qui articule les données de la sensation et les connaissances acquises déjà constituées en mémoire. Pour contourner cette difficulté, les théoriciens de la *Gestalttheorie* ont développé des expériences portant sur des figures géométriques élémentaires afin d'éliminer, selon leur expression, les « significations surajoutées par l'éducation ».

> *La théorie de la Forme (...) admet que le processus physiologique qui résulte d'un ensemble d'excitations tend à s'organiser spontanément suivant certaines lois de structure, indépendantes en principe des significations surajoutées par l'éducation. Pour étudier ces lois, la méthode la plus simple est de se donner un matériel dépourvu de signification particulière, et de le faire varier en observant, sans idée préconçue et de façon aussi naïve que possible, les « phénomènes » résultants.*
>
> Guillaume, P., 1937, p. 52.

14. Selon Delorme (1982), Von Ehrenfels continue à s'inscrire dans le concept de qualité : « il se contenta de dire qu'en plus des qualités sensibles constituées par les notes, il fallait considérer la qualité formelle dans l'explication de la perception de la forme ». L'expérience qui marque le vrai tournant est celle du « mouvement apparent », ou phénomène « phi », commentée par Wertheimer. (Se reporter à Delorme, p. 27)
15. Paul Guillaume, 1937, p. 17.
16. Voir J. Petitot, article « Forme » de l'*Encyclopaedia Universalis*, 1989, p. 712-728.
17. Olivier Sacks, 1992, p. 30.

Les gestaltistes ont montré que les formes, en tant que structures, ont des propriétés particulières. Certaines sont plus « fortes » que d'autres : elles sont plus aisément repérables, mémorisables et résistent mieux aux dégradations et aux interactions avec d'autres formes voisines ou concomitantes. Les formes les plus prégnantes se prêtent à diverses déformations : anamorphoses, surcharges, *morphing*. Enfin, certaines formes paradoxales, partiellement dégradées ou délibérément ambiguës sont la source d'illusions ou de perceptions conflictuelles particulièrement appréciées dans les arts visuels et sonores[18].

Les gestaltistes ont aussi formulé des « lois » rendant compte de l'organisation de formes qui résultent de l'assemblage d'éléments simples : la proximité, la similitude, la continuité, la clôture et le destin commun. Elles trouvent application dans la musique occidentale où la combinatoire de notes et d'unités rythmiques donne lieu à des formes temporelles répondant aux mêmes règles d'assemblage. Le lecteur impatient peut aller directement au § 3.9, page 173, où nous présentons quelques exemples sélectionnés parmi les nombreuses recherches qui sont développées sur le sujet de la musique et de la théorie de la forme.

3.2. Les formes fortes et les formes faibles

La figure 4.6 présente schématiquement quelques caractéristiques des formes fortes.

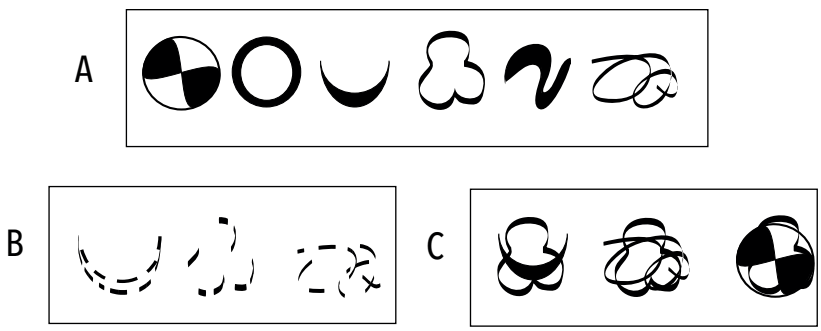


Figure 4.6 A - Six exemples de formes plus ou moins « fortes ». L'anneau, le croissant et le cercle à cadrans sont plus immédiatement repérables. B - La destruction partielle affecte différemment les formes. C - Les figures des formes « fortes » résistent mieux à la surcharge ou à la superposition.

Les formes dites « fortes » sont plutôt régulières, symétriques, avec des contours fermés, obéissant à la loi dite de « clôture ». Les formes fortes résistent plus facilement aux dégradations et aux déformations que les formes faibles et sont reconnues plus rapidement lorsqu'elles entrent en compétition avec d'autres formes. La surcharge, la destruction partielle et le masquage d'une forme par une autre permettent de tester la prégnance perceptive d'une forme.

Dans le monde naturel, on observe des formes remarquables chaque fois que la matière est soumise à des forces présentant un certain degré d'organisation spatiale ou temporelle comme l'image d'un flux s'écoulant par tourbillons alternés (figure 4.7 gauche). Par opposition, le fond d'un étang désséché offre un aspect sans organisation apparente (figure 4.7 droite).

18. Pour des exemples illustrant les propriétés des formes visuelles, voir la bibliographie (Guillaume, 1937 ; Delorme, 1982 ; Lindsay & Norman, 1980, ainsi que les sites en ligne.

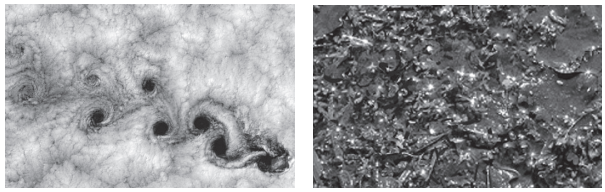


Figure 4.7 À gauche, allée de tourbillons de Karman, dans les nuages, provoquée par la rencontre du vent et une île des Juan Fernandez . À droite, image de désordre : fond d'étang désséché.

Source : À gauche, cliché Nasa, Wikipedia, allée de tourbillons de Karman.

Poussons l'analogie en passant de la matière palpable à la « matière sonore », rendue visible par l'analyse acoustique, pour tenter une typologie des formes sonores.

3.3. Peut-on parler de formes sonores ? Expérience de reconnaissance de sons très brefs

> *How sounds specify what they specify : the most important fact to keep in mind here is that EVENTS OCCUR OVER TIME, and are specified by the time course of the vibrations that stimulate the ears.*[19]
>
> Vanderveer, N. J., 1979, p. 229.

Le fait de constater que nous sommes capables de reconnaître rapidement et sûrement les sons de notre environnement implique que nous pouvons capter et extraire du signal temporel énergie/temps des données significatives sur ce qui s'est produit et que nous avons la possibilité de les interpréter. Reconnaître des sons, c'est donc reconnaître des *formes acoustiques* constituées d'une combinaison particulière de caractéristiques temporelles et spectrales, que nous avons appris à associer aux significations qu'elles portent.

Il est vrai que la notion de forme appliquée au sonore est difficile à saisir en raison même de la nature temporelle du phénomène. À propos du concept d'*objet sonore*[20], Molino remarque :

> *Le phénoménologue interroge l'œil et le visible et va jusqu'à l'invisible qui l'entoure et l'anime secrètement mais ne s'aventure jamais à interroger le sonore (...) car celui-ci semble avoir des propriétés spécifiques – inquiétantes et redoutables –, qui en font quelque chose qui n'est pas un objet comme les autres : il se déroule dans le temps, il n'est donc pas stable, posé devant nous comme une réalité tridimensionnelle dont nous pouvons faire le tour et fixer les limites.*
>
> Molino, J., 1999, p. 123.

Dès la première milliseconde et pendant toute la durée de l'émission sonore, l'activité perceptive se développe, se transforme continuellement au fur et à mesure de l'arrivée de nouvelles ondes sonores. L'interprétation finale qu'en donne un auditeur dépend donc en grande partie de la durée du son donné à entendre. Ce point est rarement pris en compte dans les expériences de reconnaissance de sons.

19. Comment les sons disent ce qu'ils disent : la donnée la plus importante à garder en mémoire est que les événements se produisent au cours du temps et que la spécificité leur est donnée par l'évolution temporelle de la vibration qui stimule les oreilles. Traduction : M. C.
20. Développé par Pierre Schaeffer (1966).

Expérience de reconnaissance de sons

Pour en prendre conscience, voici une expérience d'écoute portant sur deux collections de sons de même nombre (29), très brefs (durée égale ou inférieure à 600 ms) et totalement déconnectés de leur contexte de production. Ainsi privés de toute possibilité d'anticipation, nous ne pouvons réagir qu'aux caractéristiques acoustiques du signal sonore.

La première collection rassemble des sons dont nous avons vérifié qu'ils étaient connus des auditeurs occidentaux. Dans le montage sonore, ils s'enchaînent rapidement, sans aucune logique, ce qui produit un effet plutôt cocasse de coq-à-l'âne. À l'écoute (Son **4.4**, collection A), chacun remarquera que la reconnaissance se produit bien avant la fin de chaque son et que cette capacité est largement partagée par tous. La deuxième collection rassemble des sons familiers de même nombre et de même durée (Son **4.5**, collection B). Cette fois, l'écoute s'avère déroutante et frustrante. Bien que nous soyons toujours en pays connu, nous avons du mal à identifier les sons. Il faudrait plus de temps d'écoute, plus de silence entre chacun d'eux. Manifestement, la tâche est difficile dans le temps imparti[21].

Son 4.4 (21'')

Collection A (29 sons brefs)

Son 4.5 (19'')

Collection B (29 sons brefs)

D'où vient que nous ayons plus de difficulté à caractériser les sons de la collection B bien que les conditions d'écoute et de présentation soient les mêmes que précédemment ? Il s'agit pourtant d'événements sonores extrêmement communs qui, paradoxalement, sont souvent plus familiers que ceux de la collection A puisqu'ils sont issus d'activités ordinaires : le froissement d'une feuille de papier (5) ou d'aluminium (18), de l'eau qui bout (20), le frottement de tissus (6) ou le grincement d'une chaise (26, 28). La fréquence écologique (voir § 2.3.3) de ces sons est très supérieure à celle des sons de l'exemple précédent et, de plus, ils relèvent tous du même contexte : ce sont des bruits domestiques.

L'analyse de la structure acoustique, et donc du mode de production de ces deux collections de sons, va nous apporter quelques réponses. Les figures 4.8 et 4.9 en présentent les analyses réalisées avec les mêmes paramètres[22]. Au premier coup d'œil sur les images, on remarque une différence frappante entre les collections A et B.

La collection A présente, dans chaque case, des tracés très individualisés et très caractéristiques : ils sont tous différents. Quelques sons (2, 3, 4, 9, 15) ont une durée très brève comprise entre 100 et 200 ms. Les autres durent de 200 à 600 ms. Tous ont un début et une fin, à l'exception des sons 1, 5, 19 qui ont été extraits de séquences plus longues.

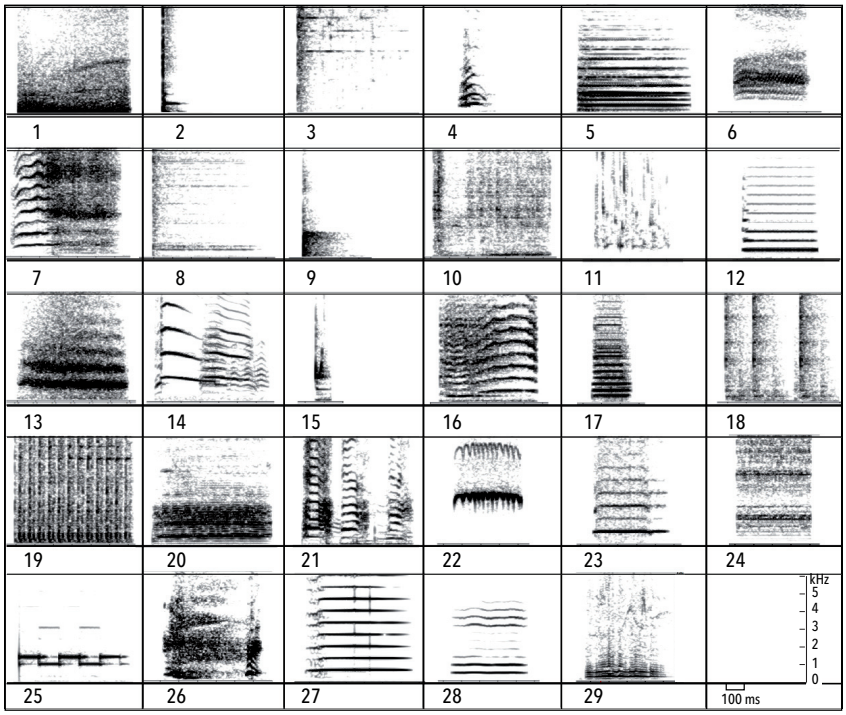
La collection B se présente plutôt comme une sorte de patchwork de trames variées (à l'exception des numéros 4 et 11), parmi lesquelles on distingue des traits verticaux réguliers (2, 10, 26, 28) ou irréguliers (5, 7, 9, 12, 17, 20, 23), des zones diffuses de points erratiques (1, 8, 13, 14, 18) ou rythmées temporellement (3, 6, 19, 22, 24), ou encore un mélange complexe de chocs, de souffles et de zones formantiques (15, 16, 24, 25, 27, 29). Plusieurs sons de la collection B produisent des images similaires.

En comparant les caractéristiques spectrales des deux collections, on remarque aussi que les sons de la collection A possèdent des fréquences bien définies – harmoniques ou non – alors que ceux de la collection B en sont dépourvus (16 et 25 exceptés). Cette spécificité est à mettre en relation avec le fait que les sons de la collection A sont rapidement identifiables.

21. Le contraste perceptif à l'écoute des deux collections est apparu fortement lors du test de catégorisation libre, présenté page 191. La durée de réalisation du test passait de 15 min (sons A) à 1 h, voire 1 h 30 pour les sons B.
22. Fréquence d'échantillonnage (fe) = 44,1 kHz ; FFT : 1 024 pts (soit 23 ms et 215 Hz) ; échelles à l'écran : 6 kHz en fréquence ; 700 ms en temps.

Son 4.4 (22″)

Collection A

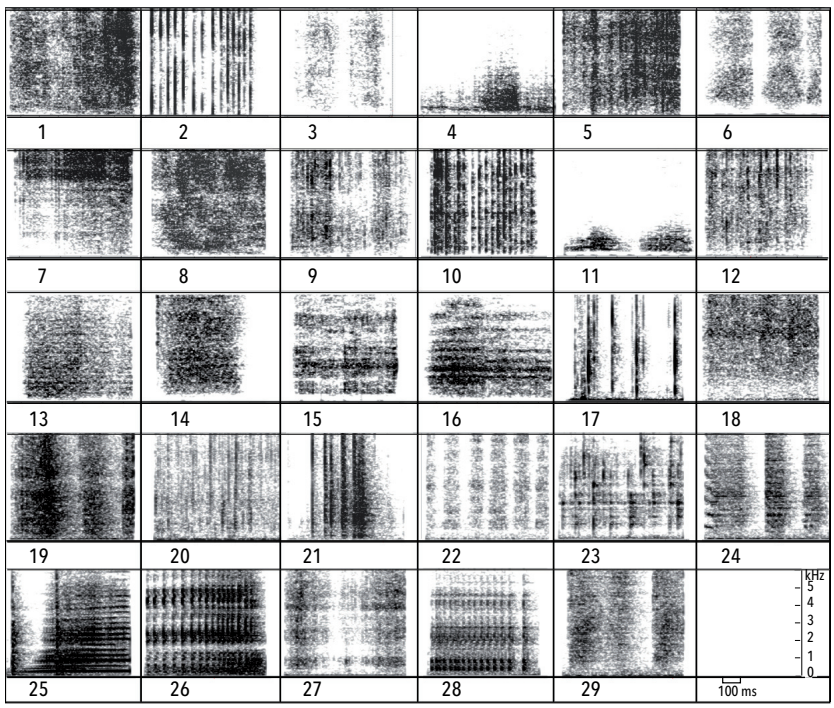


Collection A

Figure 4.8 Analyse spectrotemporelle des sons de la collection A. En abscisse : le temps (700 ms pour chaque case). En ordonnée : la fréquence en échelle linéaire (0 à 6 000 Hz). Chaque son est numéroté selon l'ordre d'apparition dans le montage sonore 4.4.

Son 4.5 (21″)

Collection B



Collection B

Figure 4.9 Analyse spectrotemporelle des sons de la collection B. En abscisse le temps (700 ms pour chaque case). En ordonnée : la fréquence en échelle linéaire (0 à 6 000 Hz). Les numéros correspondent à l'ordre d'apparition des sons dans le montage sonore 4.5.

3.4. Typologie acoustique des formes sonores

La représentation spectrale de type sonagraphique n'est qu'une convention de représentation graphique des sons parmi d'autres, cependant nous avons constaté par expérience qu'elle était particulièrement pertinente perceptivement. Elle produit des images dont les formes sont caractéristiques de la production sonore qui elle-même résulte de la combinaison du mode d'excitation (qui conditionne la *forme temporelle)* avec les propriétés mécaniques de la structure en vibration (qui conditionnent la *forme spectrale).* La forme sonore proprement dite est spectrotemporelle ou temporospectrale, selon que l'on souhaite mettre l'accent sur l'une ou l'autre de ces dimensions.

Forme temporelle : le mode d'excitation

Il n'y a que deux façons de produire des sons : soit par apport ponctuel d'énergie, c'est l'excitation par impulsion ; soit par apport continu d'énergie, ce sont le souffle et le frottement. Cette distinction se retrouve dans tous les domaines du sonore. La forme temporelle des sons impulsionnels est fondamentalement asymétrique. Elle est caractérisée par la montée rapide (1 à 2 ms) de l'énergie acoustique à son maximum, suivie d'une décroissance de durée et de pente variable. À l'opposé, les sons entretenus présentent d'importantes variations énergétiques, aussi bien au début que pendant l'entretien. Cette première distinction, qui départage les deux grandes familles de formes sonores, est portée par la forme de l'enveloppe temporelle (voir chapitre 2, page 47).

Dès les premières millisecondes, la présence d'un « transitoire » brusque donne une information non équivoque sur le mode de production du son (voir figure 4.10 a, b et c). Cette forme est si prégnante que le simple renversement temporel d'un son percussif le rend généralement méconnaissable (Son **4.6**).

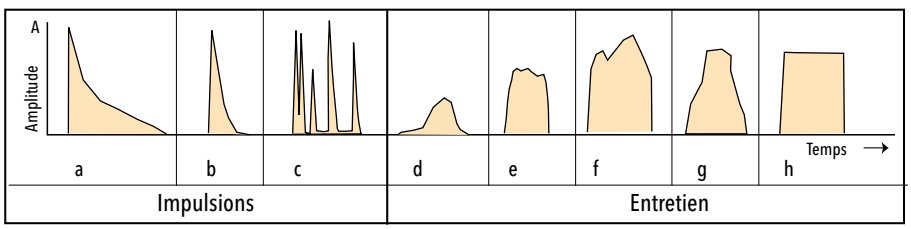


Figure 4.10 Forme temporelle de quelques exemples des deux grandes catégories de sources sonores : impulsionnelles et entretenues.

Un claquement ou une explosion imprévisibles nous font sursauter, car c'est une discontinuité temporelle majeure, une « catastrophe » au sens de R. Thom[23]. Cependant, à la différence des pigeons qui s'envolent en groupe au moindre bruit explosif, nous avons la possibilité d'interpréter la « signature » de cette impulsion en nous fondant à la fois sur la durée et sur la répartition spectrale de l'énergie dissipée, lesquelles dépendent des caractéristiques mécaniques de la structure excitée.

Forme spectrale : la réponse vibratoire de la structure excitée

Un corps ébranlé par percussion ou par frottement vibre selon des configurations modales particulières dont chacune correspond à une fréquence définie (voir chapitre 1, § 2.4). La répartition spectrale et l'évolution temporelle de ces fréquences sont caractéristiques de la structure excitée ou, plus précisément, de l'interaction entre l'excitateur et le système vibrant.

23. Thom, R., 1981. *Modèles mathématiques de la morphogenèse*. Paris, Christian Bourgois.

Deux catégories principales sont à considérer : les spectres continus (figure 4.11 b, c, d, e) et ceux qui présentent des fréquences isolées, inharmoniques (figure 4.11 a), ou harmoniques (figure 4.11 f, g, h).

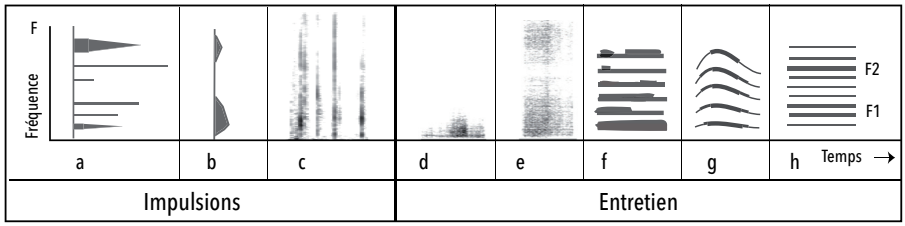


Figure 4.11 Forme spectrale de quelques exemples des deux grandes catégories de sources sonores : produites par impulsion et entretenues.

La présence de raies spectrales est caractéristique des sons de la collection A (figure 4.8) qui sont les plus aisément reconnaissables (signaux d'avertissement, sons d'instruments de musique, « voix » humaines ou animales. Un spectre de fréquences possède une structure microtemporelle tout à fait remarquable : il suffit de quelques périodes pour interpréter une sensation de hauteur, de sorte que la discrimination entre les structures excitées peut s'opérer en moins d'un dixième de seconde. Ce type de son émerge sur la plupart des bruits de fond qui sont constitués en majorité de chocs erratiques, de sons variables en intensité et en spectre. Parmi les sons périodiques, la voix humaine est reconnaissable aux formants des voyelles (voir Glossaire) qui lui sont toujours associés (voir figure 4.11 h, zones F1 et F2). Ce sont les variations spectrotemporelles de ces formants, combinées aux bruits des articulations consonantiques (chuintantes, plosives) qui constituent la matière sonore de la voix parlée humaine, reconnaissable entre toutes, quelle que soit la langue (voir le chapitre 9, page 453).

3.5. Les formes sonores : primauté de la dimension temporelle

La typologie présentée s'applique aux sons produits par des actions mécaniques usuelles, celles dont nous avons acquis une expérience personnelle à travers notre interaction gestuelle avec les éléments du monde.

Son 4.6 (8'')

Cymbale inversée

Les sons qui échappent à cette typologie suscitent des réactions diverses à la première audition. Ainsi, la première fois que des auditeurs écoutent l'enregistrement inversé d'un coup de cymbale, la surprise est totale car il n'existe pas de phénomène mécanique capable de produire un son dont l'intensité croît lentement, puis de plus en plus rapidement et qui se termine par un arrêt brutal sans la moindre résonance (Son **4.6**). C'est une source acoustique mécaniquement impossible à réaliser. Même si l'effet d'inversion, déjà pressenti par Debussy et Bartók[24], s'est banalisé avec la pratique du traitement numérique, de tels sons gardent toujours un pouvoir d'attraction du fait de leur caractère « anti-causal ».

Son 4.7 (6'')

Steel-drum

L'exemple sonore suivant (Son **4.7**) joue sur l'ambiguïté de la typologie sonore. Pour les auditeurs qui ne connaissent pas encore l'instrument qui les produit, de tels sons intriguent à plus d'un titre. Comment peut-on combiner à la fois les

24. Claude Debussy, 1905, *La Mer*, IIIe partie, mesures 38-42 ; Béla Bartók, 1937, *Musique pour cordes, percussions et célesta*, 3e mouvement, mesures 51-63. Informations communiquées par Jean-Yves Naviner.

indices d'une frappe – transitoire initial très franc et inharmonicité – et ceux d'un entretien puisque l'intensité continue à croître après le début du son ? À défaut de pouvoir leur attribuer un nom d'instrument, les auditeurs naïfs concluent généralement qu'il s'agit de sons produits par synthèse (voir Gaillard, 2000).

Le troisième exemple illustre la force de la cohérence temporelle dans la constitution des formes sonores unitaires. À l'écoute du Son **4.8**, il est clair que nous percevons deux événements distincts : un son instrumental, en l'occurrence un la_3 de violon, auquel se superpose rapidement un son pur de même hauteur, animé de fluctuations d'intensité. Or le son pur, que nous entendons comme un événement autonome, est en réalité la fréquence fondamentale de ce son de violon qui a été supprimée par filtrage (avec Audiosculpt) et réintroduite 1,26 s plus tard. Pour s'en convaincre, il faut écouter l'exemple original (Son **4.9**) et s'efforcer d'oublier la dichotomie perceptive de l'exemple précédent.

Son 4.8 (4'')
Un filtrage surprenant

Son 4.9 (4'')
Son de violon original

Ces trois exemples confirment une fois de plus l'importance de la dimension temporelle dans la constitution des formes sonores perçues au niveau acoustique. Nous allons voir qu'il en est de même au niveau syntaxique pour les séquences sonores qui s'échangent dans la communication humaine.

3.6. Les formes et le matériau sonore : deux niveaux de structuration temporelle des formes acoustiques

La typologie élémentaire que nous venons de présenter rend compte des différentes formes que prend le matériau sonore des sources. Dans la pratique quotidienne de l'écoute causale, nous avons appris à établir un lien direct entre un événement vibratoire donné et la forme acoustique du son entendu et à associer, à ces formes plus ou moins prégnantes, le nom du système producteur de son : goutte d'eau, chien, bouchon (qui saute), sifflet, etc.

Partant de la classique distinction entre forme et matière, on peut envisager une autre catégorie de formes sonores, celles qui seraient dues, non au matériau en tant que tel, mais aux transformations qu'il subit pendant le temps. Prenons comme exemple les signaux d'avertissement.

3.6.1. Les signaux d'avertissement : sources acoustiques et séquences temporelles

À l'écoute du Son **4.10**, l'idée d'avertissement est encore directement associée à la forme acoustique de la source. Il en va autrement pour les signaux du Son **4.11**. Ce que nous reconnaissons comme les pompiers, la police ou une ambulance, n'est plus la source acoustique, (le son de klaxon est commun aux trois signaux), mais la structure mélodico-rythmique caractéristique de chaque séquence. Pour chacun de ces trois exemples, la signification est codée par la forme musicale : une mélodie de deux sons à intervalle musical d'un ton pour les pompiers, d'une quarte[25] pour la police, d'une tierce majeure pour l'ambulance) associée à des durées caractéristiques, voire à un rythme (ambulance). Les deux types de signaux d'avertissement sont analysés figure 4.12.

Son 4.10 (3'')
Sifflets et sonnette de vélo

Son 4.11 (12'')
Signaux d'avertissement automobiles

25. Qui fait souffrir une oreille musicale, car elle est notablement plus petite (de 28 cents) que la quarte tempérée.

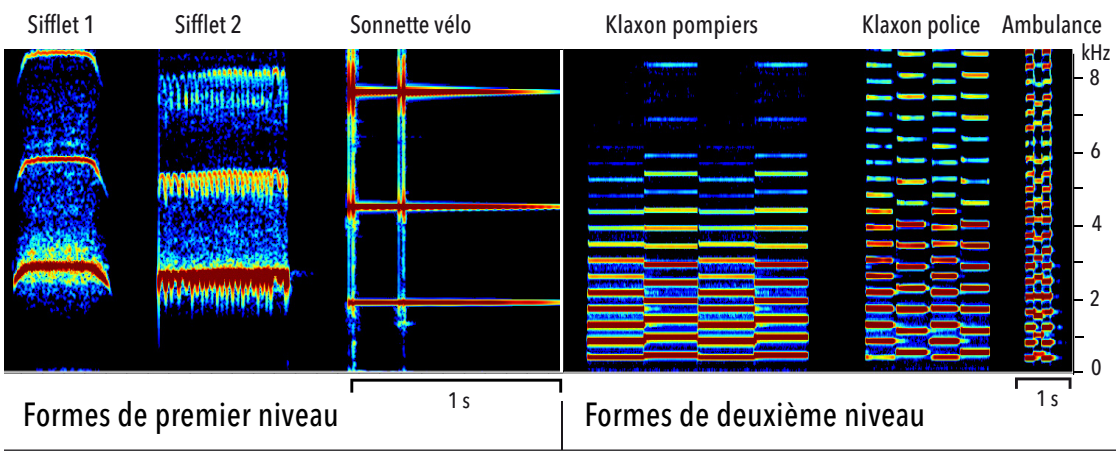


Figure 4.12 Sifflets et signaux d'avertissement. À gauche : le sifflet et la sonnette sont reconnaissables par la structure acoustique caractéristique de la source. À droite : les trois signaux d'avertissement – signal des voitures de pompiers, de police et d'ambulance – sont produits par la même source acoustique, riche en harmoniques : le klaxon. C'est la structure temporelle de la séquence mélodico-rythmique qui est reconnaissable.

D'après les sons de Vogel, C., 1999, Thèse.

Dans la partie gauche, nous voyons trois formes spectrotemporelles différentes, celles des trois sources sonores qui sont le sifflet 1 (continu), le sifflet 2 (à roulette) et la sonnette de vélo. À la partie droite de la figure, la forme acoustique de la source – le klaxon – est commune aux trois signaux. C'est un son périodique stable, riche en harmoniques, mais chaque séquence diffère par l'organisation des sons. Le décodage du sens implique une analyse des rapports entre les durées et les fréquences du son de klaxon. Nous dirons que la forme à mémoriser, la syntaxe sonore de la séquence qui porte la signification du signal d'avertissement, requiert une interprétation plus complexe que nous appellerons « de second niveau ».

3.6.2. Les formes sonores de premier et de second niveau

En résumé, nous désignerons par *formes du premier niveau* les formes acoustiques du matériau sonore lui-même, celles qui portent les traces spectrotemporelles de son mode de production et sur la base desquelles nous apprenons à discriminer les sources entre elles. Nous désignerons par *formes de second niveau* celles qui correspondent à l'organisation temporelle d'événements isolés ou aux transformations appliquées à un matériau sonore porteur. On peut les regrouper sous quatre grandes rubriques : parole, musique, signalisation acoustique et séquences sonores associées à des gestes familiers (ouvrir ou fermer une porte, se brosser les dents, poser une tasse sur une soucoupe, frotter une allumette, etc.). Les formes du second niveau, ou formes syntaxiques, sont toujours véhiculées par une matière sonore qui possède par ailleurs, sa propre forme acoustique, mais le sens qu'elles prennent est conventionnel. Il résulte d'une négociation entre les auditeurs d'un groupe donné et l'interaction « forme du matériau/forme syntaxique » du message transmis, cette dernière étant plus ou moins étroite selon les types de communication et selon les modes d'écoute des auditeurs. Même si les premières millisecondes de l'émission sonore déclenchent en priorité le décryptage immédiat de la forme du matériau sonore par l'écoute causale, celle-ci ne sollicite pas nécessairement l'attention consciente tant que le couplage source/message, généralement prévisible, se

produit. Ainsi en est-il de la voix humaine et de la parole, des sons instrumentaux et de la musique, des klaxons et des signaux d'avertissement automobiles.

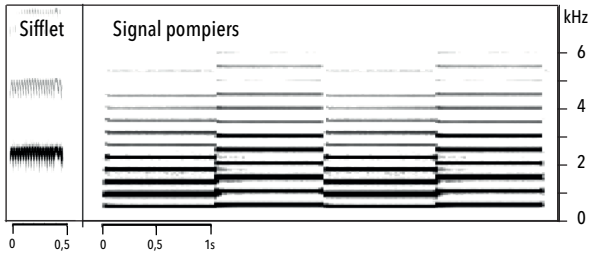


Figure 4.13 Le sifflet à roulette et la séquence *pompiers* représentés à la même échelle temporelle.

On remarquera que ces deux « niveaux » de formes – forme de la source et forme syntaxique – sollicitent très différemment l'écoute du point de vue temporel. Comme nous l'avons vu, la reconnaissance d'un son isolé s'opère dans un temps très bref, quelques dizaines de millisecondes, alors qu'une séquence constituée de plusieurs sons appelle une durée d'appréhension plus longue. On peut comparer, figure 4.13, un coup de sifflet et le signal *pompiers* analysés à la même échelle temporelle. L'alternance des deux premiers sons de klaxon, nécessaire pour l'identification de la séquence, dure quatre fois plus longtemps que le coup de sifflet.

3.6.3. Les formes acoustiques de la parole

Il est un domaine dans lequel nous pouvons faire abstraction de la source sonore[26] avec une aisance déconcertante : c'est celui de la parole. La voix humaine est éminemment changeante selon les individus et même pour un seul individu en fonction du moment.

Son 4.12 (4")

Voix (exemples 1 et 2)

Voici l'exemple (Son **4.12**) d'un mot prononcé par la même personne en voix « normale », puis en voix chuchotée.

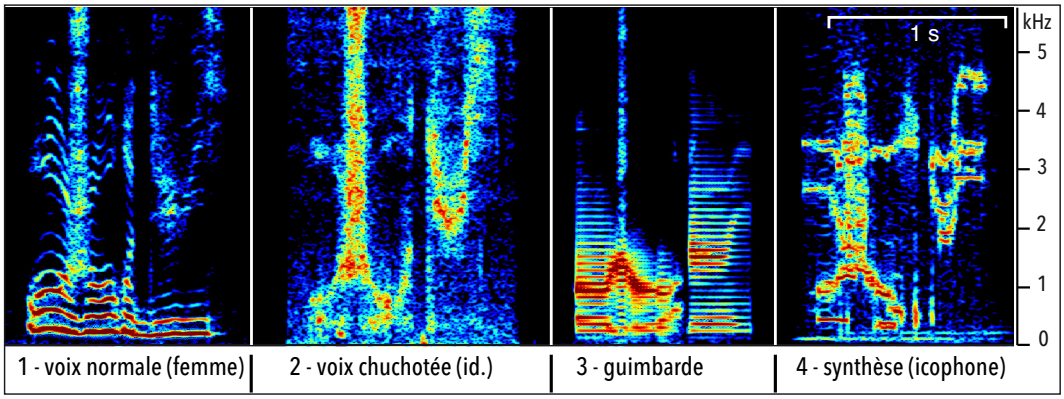


Figure 4.14 Analyse du même mot « aujourd'hui », dit avec des voix de « matière » acoustique différentes.

On voit sur la figure 4.14 que les deux signaux de parole sont très différents quant au matériau sonore : le premier est un spectre harmonique de hauteur variable (1) ; le second un bruit coloré (2). Mais la forme temporelle des variations spectrales, que nous avons appris à associer à l'articulation du mot « aujourd'hui », est bien la même sur le sonagramme.

26. Le mot « source » est pris ici dans le sens global de source sonore : le son produit par l'ensemble de l'appareil phonatoire humain (larynx et cavités de résonance).

Par analogie avec les graphies de la figure 4.15, cet exemple confirme l'existence des deux niveaux de structuration d'une forme acoustique et corrélativement, des deux modalités d'écoute déjà évoquées.

Figure 4.15 Forme et matière sonore. Analogie visuelle qui présente le même signe phonétique tracé selon des styles graphiques différents

Son 4.13 (4'')

Voix (exemples 3 et 4)

En quelque sorte, pour « entendre » la parole il ne faut pas écouter la voix qui parle, car une attention trop vive aux qualités acoustiques de la voix porteuse de parole détourne de la compréhension du sens. Le Son **4.13** donne à entendre deux exemples de dissociation entre la forme acoustique de premier niveau, le matériau sonore de la source, et la forme acoustique de deuxième niveau, constituée par les transformations spectrotemporelles de cette source. Deux voix peu ordinaires articulent le même mot.

Dans celle de la guimbarde[27] (3), le son quasi harmonique de la lame vibrante fait office de cordes vocales et l'articulation est produite par les mouvements de la langue. La « voix de l'icophone »[28] (4) est un bruit synthétisé reproduisant les formes spectrotemporelles d'une articulation humaine. Ces quatre voix prononcent le même mot.

3.6.4. Les formes de la musique

L'équivalent musical d'une variabilité de la matière sonore n'affectant pas la forme mélodico-rythmique est celui d'une même phrase musicale interprétée successivement avec trois sources musicales différentes.

Son 4.14 (17'')

Piano

Dans cet exemple proposé par Pierre Schaeffer[29], on entend successivement le piano, le chant et le marimba (sons **4.14**, **4.15**, **4.16**). Nous pourrions dire que c'est la même mélodie avec trois timbres différents : celui du piano, celui de la voix chantée, celui du marimba.

Son 4.15 (21'')

Voix

Cependant, le terme timbre est ambivalent : il désigne aussi bien la source ayant produit une matière sonore donnée que les différences de qualité du son de cette source. Nous renvoyons le lecteur au chapitre 7 dans lequel nous discutons des deux conceptualisations du timbre, soit identitaire, soit qualitative.

Son 4.16 (15'')

Marimba

La figure 4.16 montre, d'une part, la similitude du rythme (synchronisation des durées) et de la mélodie (variations concomitantes de la ligne de fréquence fondamentale) et, d'autre part, les différences spectrotemporelles qui sont dues aux caractéristiques acoustiques des trois sources : son percussif du piano et du marimba, son tenu de la voix chantée.

27. Précisons que la personne qui parle avec une guimbarde ne produit aucun son avec son larynx : sans guimbarde l'articulation est silencieuse.
28. L'icophone est un synthétiseur lecteur d'images sonagraphiques construit au LAM. Voir Leipp & coll, 1971. Il a permis de vérifier les hypothèses de la théorie de la forme en perception sonore, et en particulier en synthèse et en reconnaissance de la parole. Voir Jean-Sylvain Liénard, 1977.
29. *Solfège de l'objet sonore*, INA-GRM ; CD-1, exemples 08 et 09, « Thème rythmique » extrait de *Durboth* de Guy Reibel.

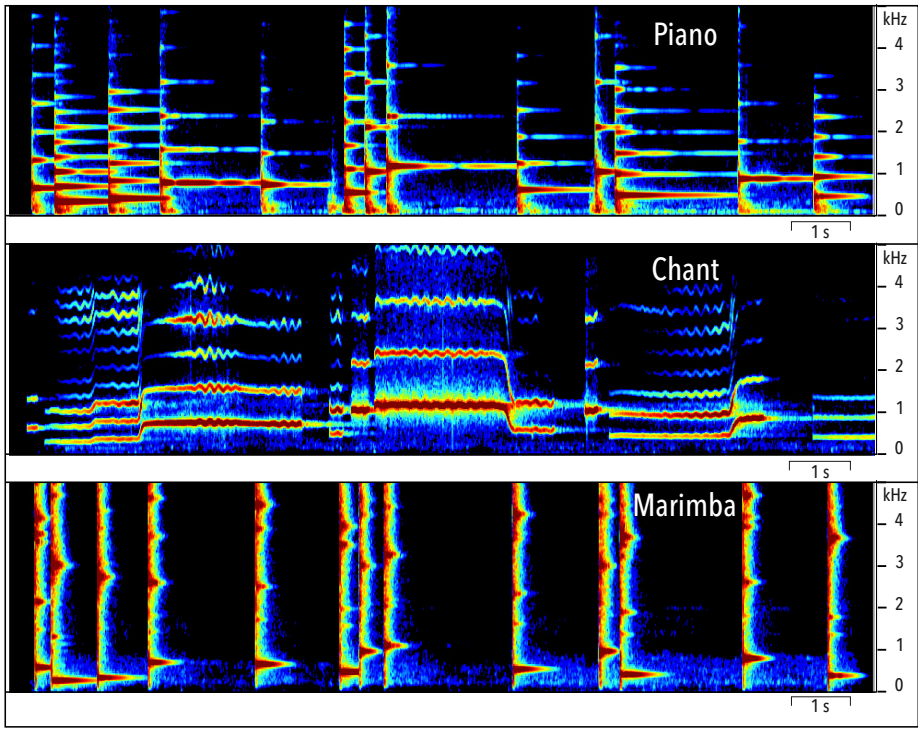


Figure 4.16 La même mélodie jouée sur deux instruments, et chantée (vocalisée). Les analyses ont été ajustées temporellement.

3.6.5. L'interprétation et la segmentation des formes de deuxième niveau (les silences)

L'exemple suivant (Son **4.17**) met en évidence un problème difficile, celui de la segmentation des séquences. En effet, c'est l'auditeur – pour qui elle fait sens – qui décide de l'unité d'une séquence acoustique de second niveau. Le début et la fin ne correspondent pas nécessairement à des indices acoustiques repérables. Nous avons sélectionné trois exemples de séquences sonores analysées figure 4.17, 1 à 3

Son 4.17 (18'')

Trois séquences sonores

- un exemple de sons liés à un enchaînement de gestes courant : ouvrir un placard dont la porte grince, prendre une bouteille et refermer la porte,
- la phrase « j'ai mangé tous les champignons » prononcée par une voix féminine,
- une courte mélodie musicale jouée à la flûte.

Les silences ont des rôles très différents dans l'interprétation de telles séquences. En parole, les silences des consonnes plosives /t/ et /p/ ne sont pas perçus comme des interruptions, car ils sont constitutifs de la forme acoustique des mots : leur suppression rendrait d'ailleurs la parole inintelligible. Dans la phrase musicale, les silences d'articulation font partie intégrante de l'interprétation du musicien. Ils varient d'une exécution à l'autre, sans affecter l'intégrité de la phrase musicale. On remarquera ensuite que tous les événements de la séquence « environnement » sont séparés par des silences. L'écoute unitaire de cette dernière se construit sur la base de la reconnaissance d'une succession de sons-sources dont l'articulation fait sens : le clac suivi d'un couinement de fréquence descendante est celui de la porte qu'on

ouvre ; les chocs musicaux caractéristiques d'un récipient, ici une bouteille, puis les événements inversés de fermeture de la porte (couinement ascendant et clac) qui clôturent la séquence. De tels enchaînements sont extrêmement courants dans notre environnement sonore quotidien, mais leur interprétation nécessite généralement de connaître le contexte et de disposer d'indices visuels.

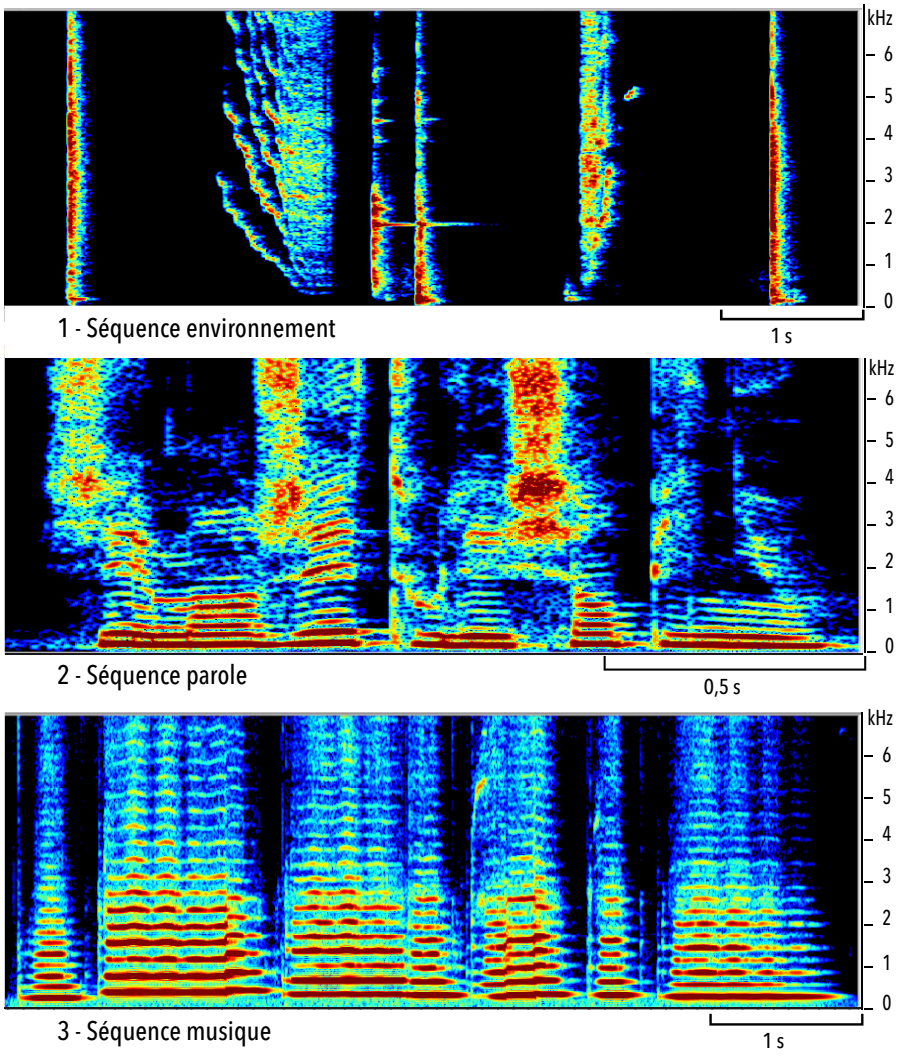


Figure 4.17 Trois séquences sonores. 1/ Environnement : séquence sonore de l'ouverture d'un placard, saisie d'une bouteille et fermeture du placard (6 s). 2/ Parole : phrase dite par une voix féminine, « J'ai mangé tous les champignons » (1,7 s). 3/ Musique : mélodie jouée à la flûte traversière (5,5 s).

Dès lors, on comprend que les séquences posent de difficiles problèmes aux ingénieurs désireux de segmenter automatiquement des chaînes continues de parole et de musique, car plusieurs niveaux de formes sont appréhendables. La structure à identifier se trouve pour partie dans le signal acoustique et, pour une partie plus grande encore, dans la mémoire des auditeurs qui ont appris à la repérer.

3.6.6. Un schéma du traitement des formes sonores

Les formes qui nous concernent en perception sonore relèvent de deux niveaux de structuration différents du signal acoustique : d'une part, les formes acoustiques des sources génératrices du son et, d'autre part, celles des séquences de sons articulés dont la syntaxe temporelle requiert un traitement spécifique à chaque sorte de communication entre humains : parole, musique, bruits et signaux d'avertissement.

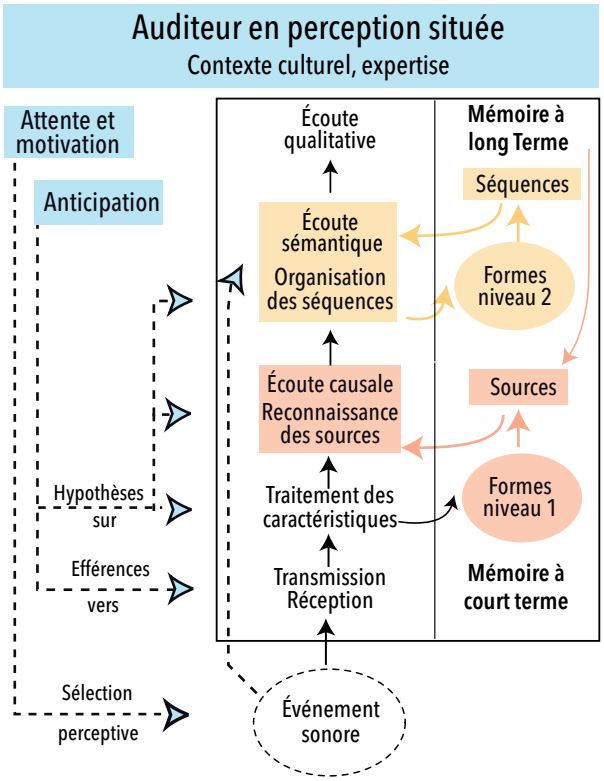


Figure 4.18 Proposition d'un schéma illustrant les principales étapes du traitement « descendant-ascendant » (ou *top-down* et *bottom-up*) d'un événement sonore.
Voir la figure 4.1, page 142

Le schéma de la figure 4.18 propose une vision des échanges entre mémoire à court terme et mémoire à long terme, en figurant les « efférences » descendantes qui interagissent lors des différentes étapes du traitement perceptif de la reconnaissance des formes. Les significations des formes pour chaque sujet sont stockées en mémoire à long terme. Projetées en amont par anticipation, elles permettent d'opérer plusieurs courts-circuits dans la partie ascendante du traitement.

Dans cette présentation, le traitement et la reconnaissance des formes, qui combine d'incessants allers-retours entre les inférences provenant des attentes d'un auditeur expérimenté et l'analyse des caractéristiques des sons, répond en partie au paradoxe souligné par quelques auteurs : « lorsqu'on sait quoi regarder, il est plus facile de voir » (Lindsay et Norman, 1980, p. 12).

3.7. Les formes sonores et la variabilité

3.7.1. La variabilité des sources

Son 4.18a (6'')

Six gouttes d'eau

Son 4.18b (6'')

Six aboiements de chiens

La notion de forme définie par des rapports entre les parties qui la constituent permet de comprendre que des signaux similaires dans leur structure, mais différents dans leurs dimensions physiques, puissent être considérés comme relevant de la même forme de base. À titre d'exemple, nous avons collecté[30] plusieurs exemplaires de deux sources sonores fort communes : la goutte d'eau et l'aboiement de chien : Sons **4.18a** et **4.18b**.

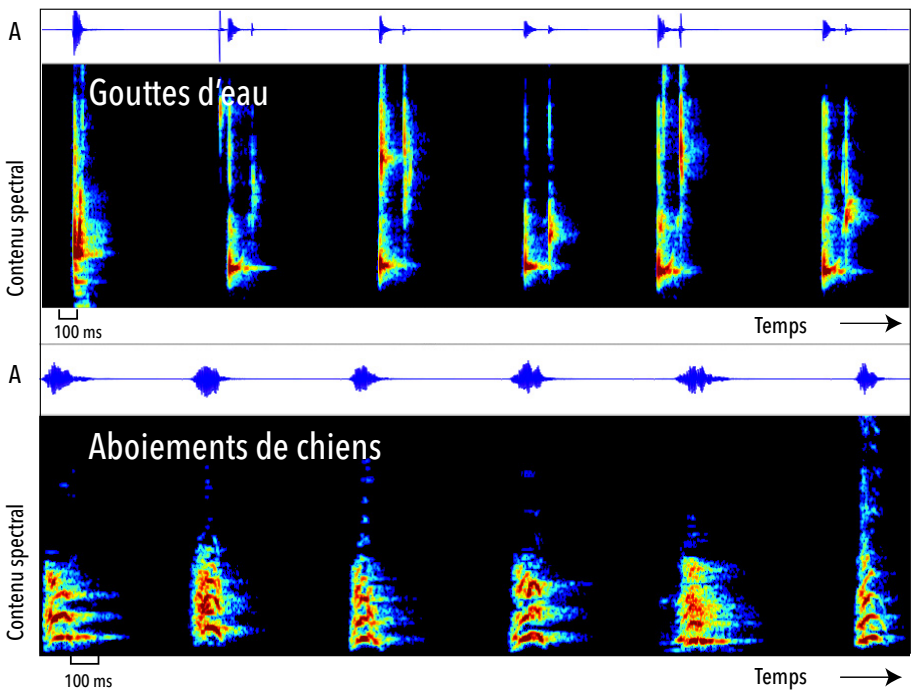


Figure 4.19 Analyses de différents sons de goutte d'eau et d'aboiement de chiens. Les formes du signal temporel (en bleu) aussi bien que les images spectrales sont caractéristiques de la source, tout en présentant une certaine variabilité d'un item à l'autre.

La représentation sonagraphique de ces deux types de signaux évoque l'écriture manuscrite (figure 4.20) ou les frises géométriques tracées à la main dont les éléments successifs sont similaires, mais jamais identiques. On peut toutefois relever des caractéristiques propres à chacune des deux sources. Par exemple, pour la goutte, un signal percussif simple ou multiple (rebonds), une composante fréquentielle principale variable, une durée brève de 100 à 200 ms. De la même façon, l'aboiement de chien peut être caractérisé de la façon suivante : un signal énergétique en forme de cloche (croissance et décroissance rapide), un spectre harmonique basse fréquence plus ou moins bruité, une hauteur fondamentale évoluant aussi en cloche, une durée brève[31] de 100 à 200 ms. Les onomatopées utilisées pour imiter ces signaux sont d'excellentes transpositions phonatoires de la signature prototypique d'un événement sonore : en français, on connaît « plic, ploc » pour la goutte d'eau et « wouh » ou « wouah » pour le chien.

30. Montage de gouttes enregistrées par l'auteur sauf la première ; montage d'aboiements de chiens de sources diverses.
31. Remarquez que la durée visible sur l'analyse est plus longue, car les enregistremnents sont faits dans des pièces sonores et non en plein air.

L'analyse de sons réels se heurte immédiatement au fait que la variabilité des formes sonores est assez considérable, surtout si la source émettrice est un être vivant. S'y ajoutent les variabilités dues aux conditions de propagation des ondes sonores et, en dernier ressort, celles dues à la position des capteurs (oreille ou microphone) dans le champ sonore.

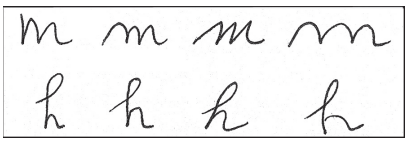


Figure 4.20 Variabilité des formes manuscrites de l'écriture des lettres « m » et « h ».

Comprendre comment nous traitons la variabilité des formes sonores nous donnerait des clés pour formaliser la typologie acoustique des instruments de musique. En effet un instrument de musique est une source sonore reconnaissable par les caractéristiques identitaires de son mode de production, offrant plusieurs possibilités de variations des qualités des sons avec lesquelles « joue » l'instrumentiste.

À l'écoute, nous pouvons tour à tour ignorer certains aspects de la variabilité lors de l'écoute identitaire, ou au contraire y porter attention lors de l'écoute qualitative.

3.7.2. Les anamorphoses sonores : parole et musique

L'anamorphose est une technique de variation des formes qui s'applique préférentiellement à certains rapports dimensionnels de la structure (figure 4.21). Au-delà d'un certain seuil de déformation, la forme n'est plus reconnaissable.

Les anamorphoses visuelles ont été largement pratiquées à la Renaissance, soit pour produire un effet grotesque, soit pour occulter ou dissimuler des formes licencieuses ou au contraire pour capter l'esprit de l'observateur intrigué, tel le crâne anamorphosé par Hans Holbein dans le tableau *Les Ambassadeurs* (1533).

Figure 4.21 Anamorphose d'un visage. La technique du quadrillage montre ici une anamorphose unidimensionnelle de la largeur de la figure selon Vignole d'Anti (milieu du XVIe siècle)

Baltrusaitis J., 1969, *Anamorphoses, ou magie artificielle des effets merveilleux*, Paris : Olivier Perrin, p. 33.

Les anamorphoses temporelles et/ou fréquentielles sont communes dans les séquences sonores de parole et de musique.

Notre premier exemple concerne l'anamorphose fréquentielle de la parole. Nous avons tous l'expérience directe des anamorphoses de la fréquence fondamentale de la voix, liées à l'âge (enfant, adulte) et au sexe (homme, femme).

Celles, plus complexes, de la position spectrale des formants vocaliques (voir le chapitre 9, § 1.4.2) est intéressante à explorer par synthèse.

Voix chuchotée synthétisée à l'icophone. Anamorphose fréquentielle en quatre étapes : 1,5 ; 1,25 ; 1 (normal) ; 0,75) (voir page 203)

Son 4.19a (12'')

Phrase complète dite par les quatre voix synthétisées

Son 4.19b (4'')

Le mot « chat » anamorphosé

Son 4.19c (4'')

Le mot « toilette » anamorphosé

Nous proposons un exemple synthétisé à l'icophone (Son **4.19**) dans lequel on entend la même phrase réalisée avec quatre anamorphoses fréquentielles différentes. À l'écoute, la voix change de qualité, pour certains, de « hauteur », mais l'intelligibilité est conservée. Or, cette synthèse simule une voix chuchotée, dépourvue d'harmoniques, donc sans hauteur fondamentale[32]. Le changement de hauteur ressenti provient de l'anamorphose fréquentielle – qui est ici considérable puisqu'elle atteint une octave – et concerne la forme spectrale, c'est-à-dire la forme de deuxième niveau. Observez sur la figure 4.22 la position en fréquence des formants spectraux *F1* et *F2* du « e » final de toilette. Le fait de modifier la position en fréquence des formants, tout en conservant leurs rapports caractéristiques permet de simuler successivement une voix de jeune enfant (à gauche sur la figure), une voix féminine, une voix masculine et celle d'un homme plutôt vieux (à droite sur la figure).

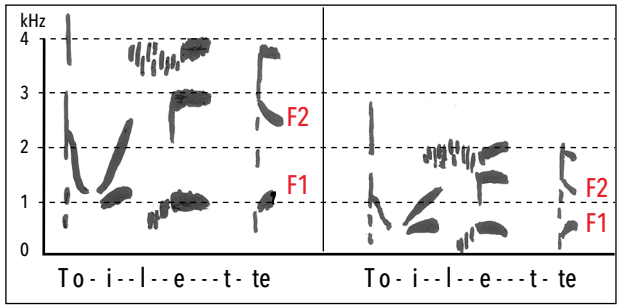


Figure 4.22 Tracé dessiné à l'icophone pour le dernier mot de la première et de la quatrième voix de l'exemple 4.19a. Pour ces deux exemples extrêmes, l'anamorphose en fréquence est dans un rapport 2. (Icophone du LAM, M. C.)

La parole naturelle est aussi continuellement affectée d'anamorphoses temporelles qui dépendent du débit général mais aussi de l'allongement de certains mots ou même de certaines voyelles qui, dans certains contextes, peuvent prendre un caractère sémantique.

En musique, les compositeurs pratiquent une grande variété d'anamorphoses en agissant aussi bien sur les rapports de durée que sur les rapports de fréquence. Les exemples de thèmes en valeurs « diminuées » ou en valeurs « longues » abondent. L'anamorphose mélodique par agrandissement des intervalles est plus rare. Citons, parmi les œuvres contemporaines, *Prologue* de Gérard Grisey[33], une pièce pour alto construite sur de multiples variations d'un thème inital par anamorphoses de durée, d'intervalles et de sonorité. De leur côté, les compositeurs de musique acousmatique, confrontés à l'irrépressible écoute causale déclenchée par les sons qu'ils enregistraient, ont redécouvert et exploité avec bonheur les anamorphoses sonores pour « détruire » la forme identitaire des sources. Dans la technique analogique, les anamorphoses temporelles (obtenues par le ralentissement de la vitesse de lecture de la bande magnétique) étaient couplées à des transpositions fréquentielles, ce qui provoquait rapidement la perte d'identité, rendait méconnaissable les sons originaux et fournissait ainsi une matière sonore nouvelle pour la composition musicale.

32. Le traitement numérique permet aujourd'hui de manipuler facilement la transposition de la fréquence fondamentale de la voix indépendamment des formes spectrales. Voir le logiciel Straight de Kawahara : <http://www.wakayama-u.ac.jp/~kawahara/STRAIGHTtrial/> ; consulté en 2013.
33. G. Grisey, *Les espaces acoustiques : Prologue* (Gérard Caussé altiste), 1977. CD Accord, 465 386-2.

3.8. Les rapports entre la forme et le fond

3.8.1. Des exemples classiques en vision

> *Il ne fait guère de doute que la distinction continu-discontinu est à la base de notre perception du monde ; c'est elle, si on veut bien y réfléchir, qui motive la distinction Ouvert-Fermé de la Topologie Générale. Elle est bien connue des Gestalt-théoriciens, qui l'ont introduite avec la distinction classique : Figure-Fond, comme des sémanticiens, qui parlent de la Forme et du Contenu...*
>
> Thom, R., 1981, p. 10.

La relation entre la forme et le fond est un autre apport important de la Gestalttheorie. La forme se définit comme une entité distincte du fond : mais comment s'opère cette « distinction » ? La question est d'autant plus délicate que les signaux qui constituent la forme et le fond sont généralement de même nature physique, comme le montre l'exemple des « colonnes » de Bahnsen (Koffka, 1935).

La figure 4.23 représente deux images, chacune constituée d'un large cadre à l'intérieur duquel on voit une alternance de bandes noires et blanches aux contours découpés. Certaines d'entre elles, symétriques par rapport à l'axe vertical, suggèrent des colonnes. Les tests perceptifs ont montré que 90 % des sujets voient des colonnes noires sur fond blanc à gauche et des colonnes blanches sur fond noir à droite. La symétrie des contours donnerait donc une cohérence à la forme colonne par rapport au fond, quelle qu'en soit la couleur. Pourtant, le fait qu'une partie des sujets produit une réponse opposée indique bien que la perception d'une forme n'est pas une extraction automatique, mais plutôt la « synthèse constructive »[34] de l'observateur qui choisit d'interpréter telle partie de la figure comme une forme, au détriment du reste de la figure qu'il considère comme fond.

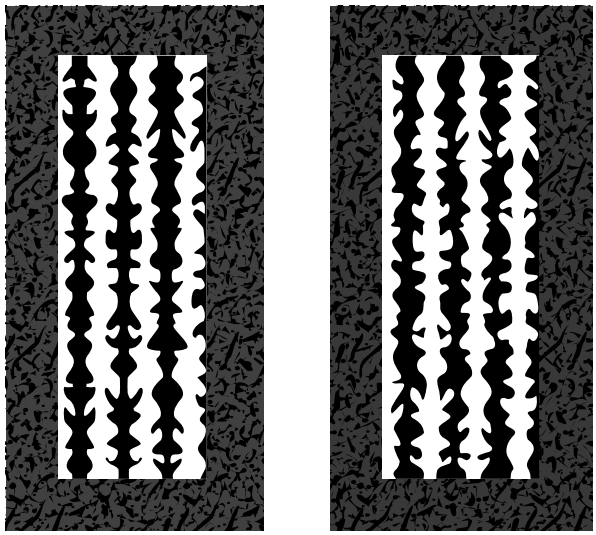


Figure 4.23 Une forme se définit toujours par rapport à un fond sur lequel elle se détache.

D'après Koffka, K.,. 1935, *Principles of Gestaltpsychology*. New York : Harcourt, p. 195.

34. « Figural synthesis » selon Neisser, 1967, *Cognitive psychology*, New York : Appleton-Century-Crofts (chapitre 4).

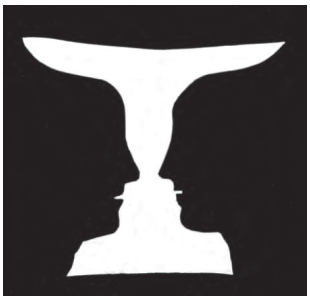


Figure 4.24 Le vase de Rubin.

Rubin E., 1921, *Visuell Wahrgenommene Figuren*, Copenhague : Gyldendhalske, figure 3 [archive.org]

Lorsque les formes sont ambiguës (tête de Boring, 1930) ou lorsque deux formes de prégnance équivalente sont imbriquées l'une dans l'autre comme pour le vase de Rubin (figure 4.24) ou pour les nombreuses créations de M. C. Escher (voir Locher, 1972), la perception oscille d'une forme à l'autre, car le choix que l'on fait d'une forme (par exemple les deux visages) exclut totalement l'autre, c'est-à-dire le vase. C'est bien le sujet qui décide à chaque instant de ce qui est la forme et de ce qui est le fond. Au besoin, il peut voir dans une figure des contours ne correspondant à aucune réalité physique, mais nécessaires à l'interprétation de la figure. Nous renvoyons le lecteur aux étonnantes images de Kanizsa (1998).

3.8.2. La forme et le fond sonores : émergences acoustique et perceptive

À la dialectique forme/fond des images correspond, dans le domaine sonore, celle de l'émergence de certains sons sur un « fond » sonore complexe. Du point de vue acoustique, un son se signale à l'attention soit parce qu'il est plus fort que les autres, soit parce qu'il se situe dans une autre zone spectrale que celles du fond sonore environnant, soit parce que sa forme spectrotemporelle est différente de celles des sons concomitants.

Du point de vue perceptif interviennent aussi la zone de plus grande sensibilité de l'oreille et prioritairement l'attente de l'auditeur qui, en anticipant intérieurement le son à venir, le détecte sur la base de quelques indices qu'il a repérés antérieurement et mémorisés.

Dans notre expérience quotidienne, nous exerçons sans cesse une activité de reconnaissance de formes sonores dans des ambiances complexes. Lorsque les sources sont extrêmement diverses et que le contexte nous permet de prévoir la plupart d'entre elles, l'écoute est aisée, voire excitante. Reprenons la scène sonore déjà entendue au paragraphe 2.4, page 149.

3.8.3. Une scène sonore complexe

Son 4.20 (16'')

Scène sonore

La figure 4.25 montre l'analyse acoustique de cette brève séquence (Son **4.20**). Quelques formes acoustiques se détachent particulièrement bien : ce sont la porte et son couinement, les voix et les rires, la musique et même la forme caractéristique du bruit du bouchon qui saute (qui est le même son que le son n° 2 de la figure 4.8, page 156). L'analyse de l'évolution temporelle des signaux met bien en évidence la spécificité de leurs formes et permet de comprendre que nous puissions aisément reconnaître des sources sonores mélangées. On remarquera aussi à l'écoute que chaque son nouveau s'impose à l'attention. S'il persiste, comme c'est le cas pour la musique, il passe au second plan et prend le statut de fond sonore dès que sa nouveauté est épuisée ou lorsque l'auditeur s'intéresse à un autre événement. Selon l'aphorisme de la théorie de l'information, « c'est ce qui change qui est intéressant ».

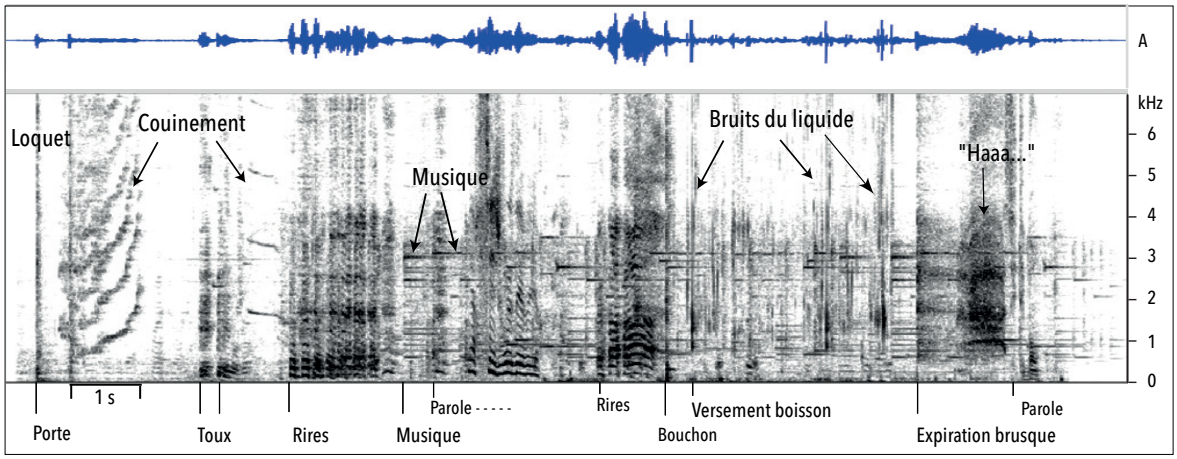


Figure 4.25 Analyse d'une scène sonore de la vie quotidienne. Les diverses sources sonores se mélangent tout en restant distinctes perceptivement et visuellement. Elles sont reconnaissables à leur forme temporospectrale.

Pour prolonger cette exploration des formes et du fond sonore, nous suggérons au lecteur de se rapporter aux figures du premier chapitre, à la figure 2.39 du chapitre 2 qui représente un paysage sonore nocturne et à la figure 2.40, page 71, qui représente l'analyse de plusieurs instruments à percussion. La notion d'émergence est reprise sous un autre angle dans le chapitre 5, § 7, page 223.

Son 4.20 (16'')
Scène sonore

3.8.4. La notion d'émergence forme-fond en musique

En musique, la diversité des relations forme-fond est à la fois plus variée et plus complexe. Les musiciens combinent à l'infini les matières sonores des instruments et des voix et créent une grande diversité de formes acoustiques. Les oppositions de sonorité, les fusions, les masquages partiels, qui font partie des ressources expressives de la musique, jouent un rôle majeur dans les musiques traditionnelles. De surcroît, pour une musique donnée, les relations forme-fond peuvent se décliner à différents niveaux hiérarchiques de la structure temporelle de la pièce musicale. Du point de vue perceptif, il faut aborder séparément les musiques combinant des hauteurs tonales, celles dont la structure s'articule principalement sur les rythmes, celles qui sont constituées de matière sonore enregistrée ou de synthèse et faire une place particulière à celles qui incluent la voix humaine. La question est trop vaste pour trouver place dans cet ouvrage et, malheureusement, rares sont les études d'analyse musicale qui prennent en compte le matériau sonore de la musique, à l'exception de ceux de Robert Erickson (*Sound structure in music*, 1975) et Robert Cogan (*New images of musical sound*, 1984) auxquels nous renvoyons le lecteur. Nous ne pouvons donner ici que deux exemples.

Exemple 1 – Musique de percussions. Dans une polyphonie de percussions, chaque instrument a généralement une signature spectrale bien spécifique, parfaitement identifiable sur un sonagramme. Voici l'exemple d'une pièce polyphonique vietnamienne[35] d'une grande complexité rythmique, que nous avons analysée à la demande de l'ethnomusicologue Trân Van Khe, afin de lever quelques difficultés de transcription musicale (Son **4.21**).

35. Trân Van Khe, « L'utilisation du sonographe dans l'étude du rythme », *Revue de Musicologie*, tome LIV, n° 2 (p. 222-232), 1968.

L'ensemble comprend cinq instruments de percussion : le double tambour de cérémonie (mâle et femelles), le tambour en sablier, la corne de buffle évidée, le petit gong, les cymbales (figure 4.26).

Son 4.21a (9'')

Trong Lay ; extrait court correspondant à l'analyse de la figure 4.26.

Son 4.21b (12'')

Montage pour écoute du petit tambour en sablier (voir page 202)

Son 4.21c (1'21'')

La pièce entière

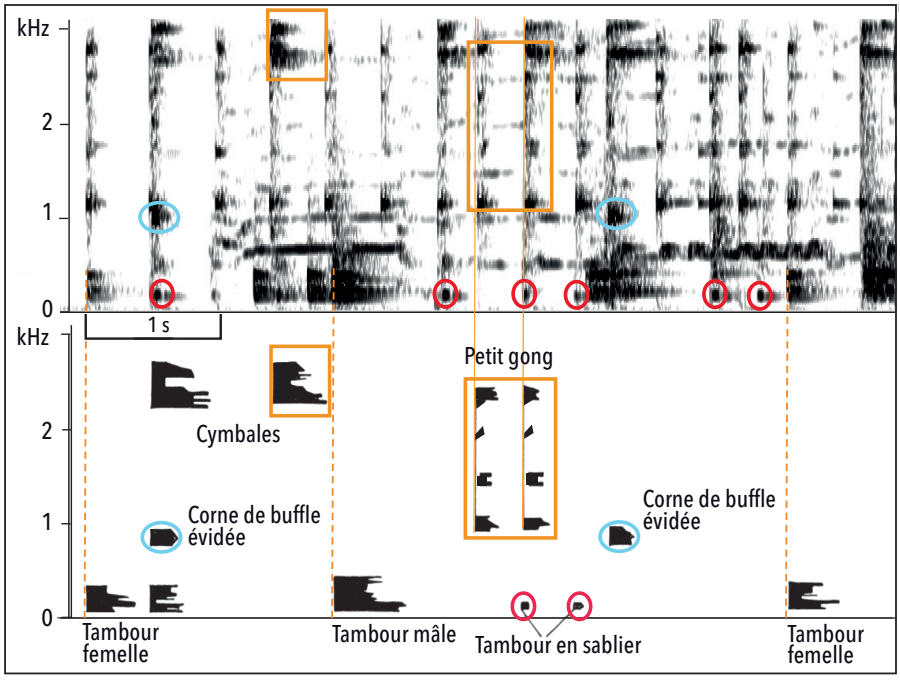


Figure 4.26 Analyse spectrale du début de la polyphonie rythmique du Viêtnam : *Trông Lay*. Notez que le hautbois (visible sur le sonagramme) n'est pas pris en compte dans l'analyse. Voir Leipp et Trân Van Khe, 1967, *Bulletin du GAM* n° 29, pour l'analyse complète de cette pièce.

La polyphonie de cette pièce est si dense que le « tambour en sablier », joué simultanément avec le petit gong ou avec les cymbales, est resté mystérieusement introuvable sur la représentation sonagraphique lors de l'analyse acoustique. Ce n'est qu'après un travail d'écoute attentive et surtout grâce à l'imitation vocale qu'en a faite l'ethnomusicologue, que nous avons pu repérer à l'oreille la note spectrale de cet instrument et, à l'aide de diverses manipulations acoustiques, rechercher le son correspondant dans la zone des 200 Hz où il se trouvait effectivement ! Pour apprécier les difficultés d'écoute posées par ce discret petit tambour, écoutez le montage sonore **4.21b** dans lequel vous pourrez l'entendre isolément, extrait de l'ensemble, puis résinséré. Il faut dire, à l'intention des ethnomusicologues, que le dépouillement aurait été grandement facilité si nous avions pu disposer au préalable de l'enregistrement de chaque instrument joué séparément.

Exemple 2 – Musique mélodique et voix. L'émergence sonore des sources d'une musique mélodique est une étude encore plus complexe, car le décodage de la perception des hauteurs, qui est très mal connu pour la polyphonie, interfère avec celui des formes spectrales des sources : on y rencontre toutes les variétés possibles de combinaisons et/ou d'oppositions entre formes mélodiques et formes spectrales.

Parmi toutes les sources musicales, la voix humaine émerge de façon remarquable. Même sans parole articulée, la voix est déjà reconnaissable par les caractéristiques vocaliques de son spectre et par les changements qui dépendent de la tessiture. Mais c'est surtout l'instabilité de la hauteur, transformée en vibrato par le travail des chanteurs, qui procure à la voix humaine une émergence acoustique et cognitive.

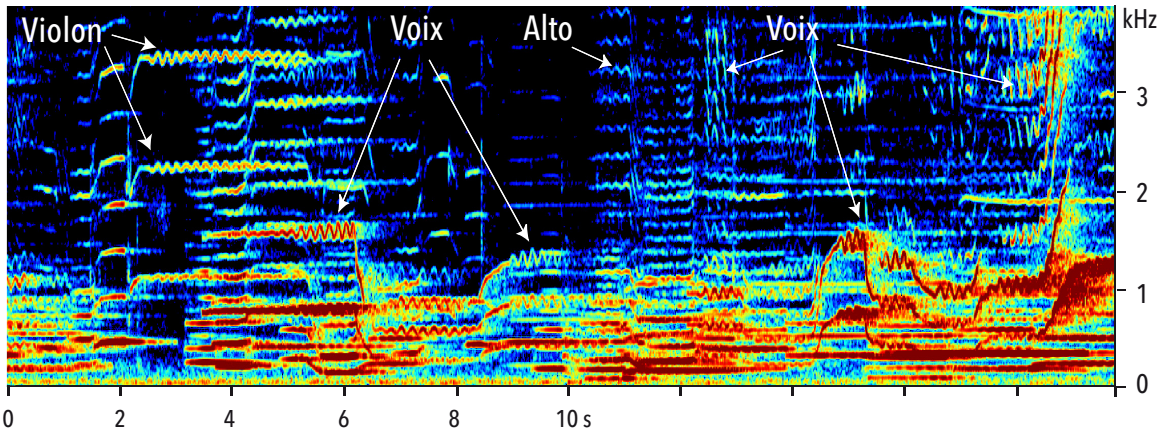


Figure 4.27 Fragment (20 s) de *Quatuor II* de Betsy Jolas pour soprano colorature, violon, alto, violoncelle (chant : Mady Mesplé). Ce fragment correspond aux mesures des pages 48-49 de la partition musicale.

L'écoute du *Quatuor II* de Betsy Jolas (1964), dans lequel la voix prend la place du premier violon, est remarquable à cet égard (Son **4.22**). Malgré les prouesses vocales quasi instrumentales de la cantatrice Mady Mesplé, la voix vibrée se situe dans un univers différent de celui des instruments à cordes, principalement en raison de l'ampleur du vibrato (comparez sur la figure 4.27, à fréquence équivalente, le vibrato vocal et celui des instruments). Enfin, lorsque s'y joint la parole articulée, de nouvelles formes émergent par leur singularité spectrale, car aucun instrument mécanique – à l'exception de la sourdine « wouah-wouah » de la trompette ou d'effets similaires à l'harmonica – ne peut produire les variations formantiques du conduit vocal.

Son 4.22 (22'')

Quatuor avec voix chantée

3.9. La musique et la théorie de la forme : l'organisation perceptive

La musique occidentale s'est développée en créant une notation des éléments de son vocabulaire. Au niveau de la syntaxe, il n'existe plus que des éléments symboliques, les notes[36], avec lesquelles les compositeurs et les improvisateurs créent des formes, dont certaines d'entre elles n'ont plus aucun lien avec la matière sonore. Notes et rythmes sont des unités que les musiciens combinent, permutent, associent selon les « lois » énoncées par les théoriciens de la Gestalt (voir page 152), en créant des formes symboliques de second niveau.

3.9.1. L'organisation monodique : sons successifs de même hauteur sonore

La figure 4.28 présente en vis-à-vis différentes organisations de signes visuels (a, b, c, d) et de signes musicaux (A, B, C, D). Une suite d'éléments alignés (à gauche) correspond à une séquence de sons identiques se produisant à intervalles réguliers (à droite). En jouant sur la similitude et sur la proximité des regroupements, on peut

36. Conduisant à « l'abstraction notale », selon l'expression de Robert Francès (p. 27).

mettre en évidence différentes structures temporelles. Ainsi en est-il des variations d'intensité (B), de timbre (C), d'espacements temporels (D,) qui forment des organisations dans des domaines différents. Avec cet exemple simple, on entrevoit déjà la possibilité d'une hiérarchisation des différents niveaux d'organisation des durées : la pulsation (A), la mesure (B et C) et les rythmes (D).

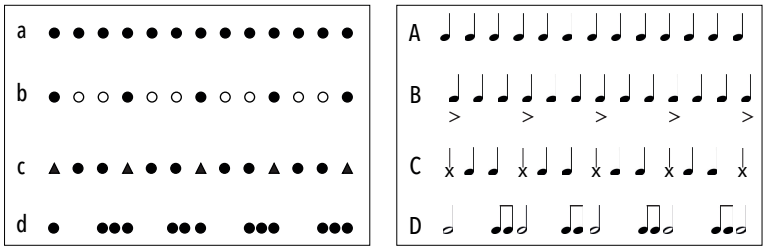


Figure 4.28 À gauche, structuration d'une séquence d'éléments figuratifs (d'après Wertheimer) ; à droite, traduction sonore et musicale en accent, timbre, durées. (M. C.)

Son 4.23 (24'')

Structuration spontanée de variations aléatoires (voir détails page 203)

Le besoin de structurer perceptivement une suite régulière d'événements semble une donnée fondamentale de la perception humaine, comme en témoigne le Son **4.23** réalisé en 1992 par Gérald Bennet. Dans cet exemple, un son est répété à intervalles réguliers (cadence environ 5/s) avec la même intensité et de petites variations du spectre qui se produisent de façon aléatoire. Pourtant, la majorité des auditeurs « entend » spontanément des groupements temporels de trois à cinq sons ou plus, donc éprouve le besoin de structurer le déroulement du temps selon un pseudo-rythme qui peut changer d'une écoute à l'autre.

Le domaine de l'organisation auditive a fait l'objet de nombreuses recherches expérimentales avec des sons de synthèse. Citons, parmi les auteurs, Leo van Noorden (1975), Albert Bregman (1990), David Wessel (1978) et Diana Deutsch (1982, 1994) auxquels nous empruntons plusieurs exemples.

3.9.2. La ségrégation mélodique : le rôle de la grandeur des intervalles et de la distance temporelle

Voici un exemple extrait de la *1re Fantaisie* de Telemann pour flûte seule, dans lequel une suite de notes peut donner lieu à l'écoute de deux mélodies entrelacées (voir la figure 4.29).

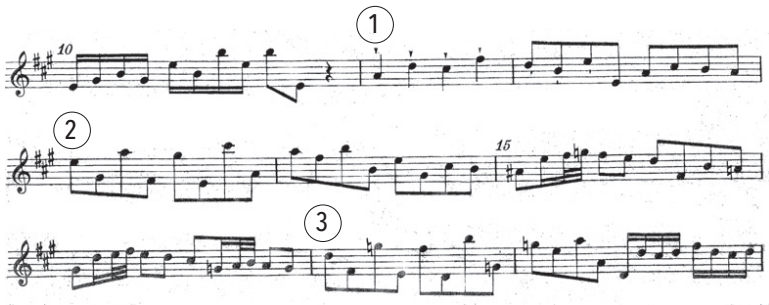


Figure 4.29 Exemple de musique instrumentale monodique à effet polyphonique.

Telemann G., *1re Fantaisie pour flûte seule* [Imslp].

Le thème est énoncé mesure 11 en *la* majeur (1). Il est repris mesure 13, en *mi* majeur (2), en alternance avec une deuxième mélodie jouant le rôle d'une basse, puis une troisième fois mesure 17 (3), en *ré* majeur. Il est intéressant de remarquer que l'apparition de deux flux se produit sur la base de la cohérence mélodique. Les intervalles entre croches successives deviennent très grands, alors que ceux des croches prises deux à deux donnent lieu à l'audition de deux mélodies évoluant dans des tessitures différentes. L'effet cesse à la mesure suivante. Ce procédé est bien connu des compositeurs de musique classique qui en usent dans l'écriture des instruments monodiques afin de produire l'effet d'une polyphonie mélodique..

Ségrégation mélodique ; expériences de Bregman *(voir page 203)*

Les relations de proximité entre sons successifs dépendent de deux sortes de distances : une distance mélodique et une distance temporelle. Voici un petit exemple emprunté à Albert Bregman (Son **4.24**), qui montre l'interdépendance du tempo et de l'écart intervallique dans la perception d'une seule ou de deux mélodies. L'exemple joue à la fois sur la tessiture et sur le tempo pour qu'une séquence de sons produise l'effet de deux mélodies indépendantes. La forte cohérence du rythme dit du « galop »[37] facilite vraisemblablement la cohésion mélodique perçue.

Son 4.24a (26″)

Deux sons séparés par un grand intervalle

Dans son travail de thèse, Van Noorden a exploré systématiquement les effets de ségrégation ou de fission mélodique. L'exemple suivant (Son **4.25**), assez complexe, met en jeu les relations de distance spectrale, de similarité rythmique et mélodique, ainsi que la loi dite de « bonne continuation » au moment du croisement. Du point de vue acoustique, la séquence n'est composée, comme précédemment, que d'une suite de sons successifs détachés, combinant deux flux réguliers de sons brefs dont les cadences, qui sont dans un rapport double l'une de l'autre, s'inscrivent dans une pulsation stricte représentée figure 4.30 (gauche).

Son 4.24b (26″)

Deux sons séparés par un petit intervalle

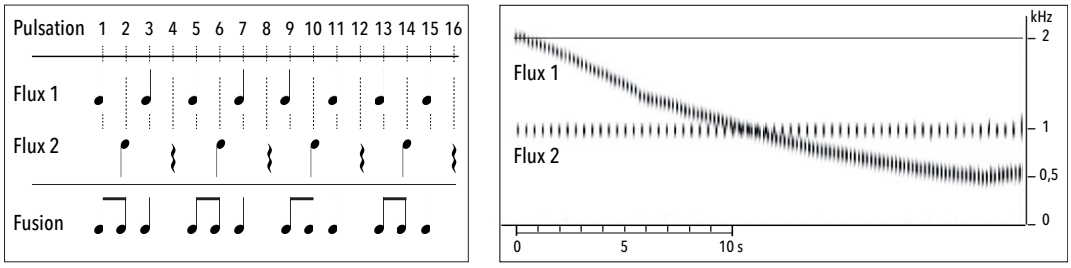


Figure 4.30 Expérience de fission-fusion mélodique de Van Noorden (1975).
À gauche, représentation rythmique des séquences mélodiques. La durée de la pulsation est de 100 ms.
À droite, analyse sonagraphique de la première partie de l'exemple sonore. Le trajet du son du flux 1 est de deux octaves (de 2 000 Hz à 500 Hz).

D'une façon générale, tous les auditeurs perçoivent d'importants changements perceptifs au cours de l'écoute. Le flux 1 suit une courbe descendante-ascendante alors que le flux 2 est fixe. À l'approche du croisement, les notes de la mélodie supérieure s'intègrent dans celles de la mélodie fixe, et soudainement la perception rythmique se transforme. Du fait de l'entrelacement des sons et de leur distribution temporelle, trois groupements sont possibles : ceux de la mélodie fixe (flux 2) à la cadence d'une blanche, ceux de la mélodie variable (flux 1) à la cadence d'une noire, puis, au moment de la fusion, le rythme croche-croche-noire (galop). Les instants de basculement changent selon les auditeurs et le rythme croche-croche-noire peut être compatible avec la perception de deux mélodies séparées. En écoutant attentivement cet exemple, on remarque que les effets perceptifs qui se produisent

Son 4.25 (45″)

Ségrégation mélodique; expérience de Van Noorden

37. On peut le traduire par le motif [croche-croche-noire] (voir figure 4.30 gauche), ou encore par l'onomatopée « ta ga da, ta ga da ».

pendant la descente de la ligne mélodique du flux 1 ne sont pas simplement inversés lors de la remontée. Une des explications tient sans doute au fait qu'en musique la ligne mélodique supérieure bénéficie d'un statut perceptif particulier. On notera aussi qu'à des changements physiques graduels correspondent des changements perceptifs brusques, indice d'une perception catégorielle (voir page 180).

3.9.3. La ségrégation mélodique : le rôle du timbre

Son 4.26 (6'')

Mélodie « arlequin »

L'alternance rapide de timbres instrumentaux différents peut perturber l'écoute d'une mélodie comme dans l'exemple sonore **4.26** qui a été réalisé par un montage en « arlequin » de différents instruments de musique. À l'appui de cette observation, l'orchestration de l'*Offrande musicale* de J.-S. Bach par A. Webern est souvent donnée en exemple. L'effet n'est pas aussi saisissant que dans l'exemple précédent car, d'une part, nous connaissons bien le thème, d'autre part, les artistes atténuent les transitions par le phrasé et surtout les transitions instrumentales se produisent à des points d'articulation syntaxiques prévisibles (Son **4.27**).

Son 4.27 (30'')

Début de l'Offrande musicale (Bach-Webern)

Expérience de D. Wessel (voir page 204)

Appliquée à une suite de sons répétés dont la cohérence mélodique est faible, l'alternance de timbre provoque, dans l'expérience de David Wessel, une fission mélodique doublée d'un effet de ralentissement rythmique analogue à l'exemple précédent de Van Noorden. On entend tout d'abord une mélodie de trois sons répétés avec le même timbre (Son **4.28a**). Dans les deux sons suivants (Son **4.28b** et **4.28c**), deux instruments différents jouent alternativement une note sur deux. La continuité perceptive bascule dans le domaine spectral et l'attention est attirée par deux groupes de trois sons descendants, entrelacés, qui se poursuivent à un tempo deux fois plus lent. Voir figure 4.31 l'alternance des spectres. Notes bleues : H1 faible, H4 dominant. Notes rouges : H1 et H2 dominants.

Son 4.28a (6'')

Une source instrumentale (figure 4.31 haut)

Son 4.28b (8'')

Alternance de deux sources

Son 4.28c (8'')

Autre (figure 4.31 bas)

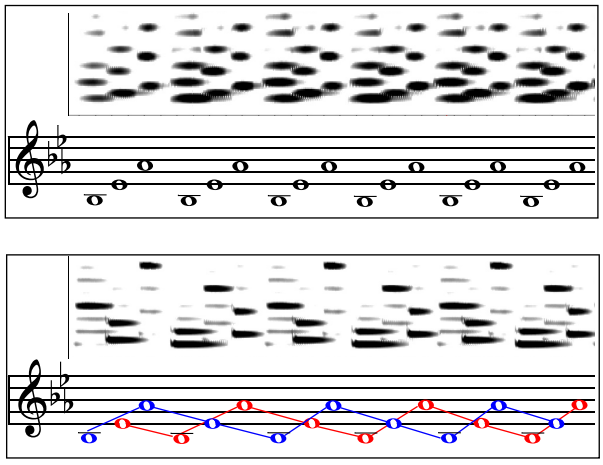


Figure 4.31 Expérience de fission mélodique par le timbre. Analyse spectrographique et notation musicale des Sons 4.28a (en haut) et 4.28c (en bas).

3.9.4. L'organisation polyphonique : le croisement de parties

En polyphonie, plusieurs mélodies évoluent de façon indépendante. Le suivi d'écoute de deux lignes mélodiques simultanées (cas le plus simple) n'est pas chose aisée : il s'apparente à l'écoute de deux conversations indépendantes.

La tâche est facilitée lorsque les deux parties sont différenciables par le timbre, la tessiture et la ligne mélodique, ce qui constitue les fondements de l'écriture polyphonique. Mais, à la différence de la parole, la musique ne convoie pas de message sémantique et recherche au contraire des situations d'ambiguïté perceptive. Comme l'ont laissé entrevoir les expériences de Van Noorden (Son **4.25**) et de David Wessel (Son **4.28**), la situation de croisement, par exemple, donne lieu à des bifurcations parfois inattendues.

Soient deux trajets mélodiques, l'un ascendant et l'autre descendant (figure 4.32). Quelle est la loi de bonne continuation mélodique au moment du croisement ? Écoutons le Son **4.29**. Il y a fort à parier que les auditeurs non prévenus – c'est-à-dire ceux qui ne connaissent pas la partition – n'auront pas entendu le croisement des voix.

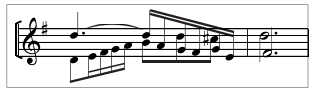


Figure 4.32 J.-S. Bach, croisement de deux voix dans la Variation Goldberg n° 3, mesure 8.

En nous reportant à la figure 4.33, nous constatons que c'est plutôt la figure B qui correspond à leur perception sonore, alors qu'en modalité visuelle la loi de bonne continuation oriente plutôt sur la figure C. Tentons une interprétation.

Son 4.29 (4'')

Clavecin 1 (diapason, $la_3 = 415$)

Pour dessiner un « X » nous procédons à l'intersection de deux droites tracées *successivement*. Pour réaliser un « X sonore », il faut deux sources interférant *simultanément*. Chaque partie ayant commencé son trajet, il semble que l'attente perceptive favorise le fait que chaque partie reste dans sa tessiture de départ. Effectivement, nous entendons assez spontanément deux mouvements contraires, deux « V » inversés. Ce phénomène a été étudié exhaustivement par Diana Deutsch[38] qui a montré de surcroît que, en situation d'écoute au casque, la partie aiguë du « V » était préférentiellement localisée du côté de l'oreille directrice[39], quelles que soient les conditions de distribution spatiale des sons. Dans ce cas, la localisation auditive serait donc indépendante des indices sensoriels.

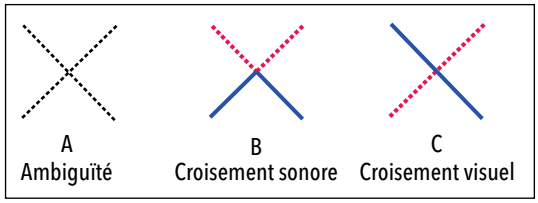


Figure 4.33 La loi de « bonne continuation » ne suit pas les mêmes règles en perception sonore et en perception visuelle.

Pour favoriser l'autonomie de deux parties, l'interprète ne dispose, sur un clavecin, que des subtilités d'articulation par les silences et les infimes décalages temporels entre sons simultanés afin de donner à chaque ligne mélodique son individualité. Un montage sonore séparant les deux parties, l'une sur la voie gauche et l'autre sur la voie droite, procure incontestablement une dichotomie mélodique (Son **4.30**). Nous y revenons dans le paragraphe suivant.

Son 4.30 (4'')

Clavecin 2 (à écouter au casque)

38. Diana Deutsch, 1994, *La perception des structures musicales*, p.115-144 ; Deutsch, 1999, *Grouping mechanisms in music*, p. 299-348.
39. L'oreille droite pour les droitiers et inversement pour les gauchers.

Dans le même ordre d'idées l'exemple, du finale de la *Sixième Symphonie* de Tchaikovsky est particulièrement surprenant, car il donne à entendre une mélodie qui n'est pas écrite explicitement sur la partition. (voir Deutsch, 1999, p. 325)

Expérience à écouter sur enceintes

Écoutons successivement la partie des premiers violons (Son **4.31a**), puis celle des seconds violons (Son **4.31b**). Lorsqu'on les écoute simultanément à l'air libre et non au casque (Son **4.31c**), une surprise nous attend : une nouvelle mélodie a surgi.

Son 4.31a (15'')

Mélodie des violons I

Son 4.31b (15'')

Mélodie des violons II

Les deux parties de violon, pratiquement homorythmiques, progressent en mouvements contraires et se croisent à chaque nouvelle note (figure 4.34).

Son 4.31c (15'')

Ensemble : violons I + II

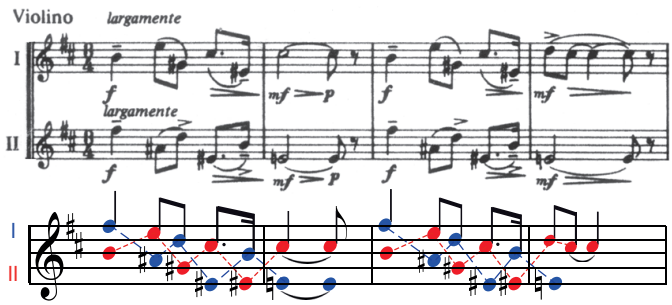


Figure 4.34 Tchaikovsky, *Symphonie n° 6*, début du finale et schéma mélodique des deux premières mesures. Du point de vue perceptif, c'est la ligne des sons supérieurs qui domine, en alternant violons I et II.

Partition musicale : site Internet Imslp.

Son 4.31d (26'')

Extrait d'orchestre

Tchaikovsky (détails page 204)

C'est sur la note la plus aiguë que s'organise la mélodie perçue qui est : *fa*#, *mi*, *ré*, *do*#, *si do*#, etc. L'effet de fusion est favorisé par la proximité des timbres, mais il serait intéressant d'entendre cette musique avec la disposition orchestrale qui se pratiquait à l'époque (Meyer, 1993), c'est-à-dire les violons I à gauche et les violons II à droite du chef d'orchestre, car la séparation spatiale des sources sonores risque alors d'entrer en compétition avec le suivi cognitif de la ligne supérieure.

3.9.5. La ségrégation des flux polyphoniques par la localisation spatiale

Expérience de Bregman à écouter sur enceintes

La localisation spatiale est un paramètre important de la séparation des sources similaires entremêlées. Il est bien connu que l'effet « cocktail party » rend difficile l'écoute de plusieurs conversations enregistrées en monophonie, pour peu que les voix des interlocuteurs soient similaires. L'expérience de Bregman (1995) atteste de l'importance de la localisation spatiale pour le suivi de deux flux entremêlés. Elle porte sur l'interprétation d'une musique de xylophones dont on n'entend d'abord que la mélodie du premier instrument. Dès qu'entre le deuxième, qui joue en alternance avec le premier, les deux mélodies, qui sont isorythmiques et se déroulent dans la même tessiture, s'entremêlent perceptivement (Son **4.32a**). Même avec une grande attention et plusieurs écoutes, il est extrêmement difficile de suivre une partie isolément tant que les deux voix sont mélangées dans chaque canal. Pourtant dès qu'elles sont distribuées, chacune sur un canal différent, la séparation s'opère aisément et nous entendons deux instrumentistes indépendants : l'un à gauche, l'autre à droite (Son **4.32b**). Cette expérience est impressionnante lors d'une écoute au casque. Il faut aussi la pratiquer en champ libre, avec deux enceintes, pour percevoir la transformation du champ sonore qu'opère le passage progressif d'un signal monophonique à un signal bi-voies.

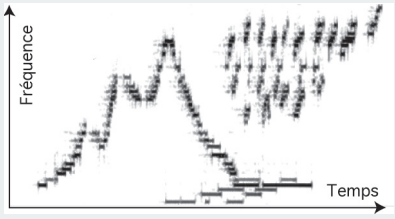
Son 4.32a (51'')

Musique monophonique (deux mélodies)

Son 4.32b (34'')

Passage progressif à une diffusion bicanale

Des formes visuelles et des formes sonores



« Une montagne avec de la pluie» : dessin à écouter.

Figure 4.35 Analyse sonagraphique des dessins peints sur un support transparent, puis « lus » par un appareil de synthèse sonore : l'icophone. En ordonnée : la fréquence ; en abscisse : le temps. (LAM, 1969)

Son 4.33a (5'')

Montagne + pluie

Le dessin ci-dessus évoque « une montagne avec de la pluie ». Tracé dans le plan fréquence/temps et lu par le synthétiseur « icophone »[1], il produit un son ascendant par paliers, puis descendant, et se termine par un fourmillement de glissandi aigus (Son **4.33a**).

Nous invitons maintenant le lecteur à écouter les sons **4.33b** à **4.33e**. Chaque brève séquence est la traduction sonore, la « sonification » de l'un des quatre dessins ci-dessous. L'axe horizontal représente le temps et l'axe vertical la fréquence. Sachant qu'une barre verticale produit un bruit impulsionnel et une barre horizontale une fréquence pure, qu'un cercle génère deux mélodies simultanées, l'une ascendante et l'autre descendante, il s'agit de trouver le dessin correspondant au son entendu.

L'écoute demande un petit effort d'attention, mais très rapidement la correspondance image-son fonctionne (réponse page 205). Ce qui surprend dans ce jeu est l'aisance avec laquelle le son **4.33d** sort du lot. Il est aisément entendu comme un mot, alors que le dessin-support, qui en est très exactement la forme acoustique spectrotemporelle, n'évoque rien de particulier[2].

Le fait que tous les auditeurs disent avoir entendu une « voix » témoigne de la force de l'association que nous avons établie antérieurement entre voix humaine et parole, même si la « matière » acoustique de cette voix est ici bien différente de celle d'un humain.

Cette expérience met aussi l'accent sur les difficultés rencontrées dans l'interprétation des analyses acoustiques. Si le plan sonagraphique offre une représentation privilégiée pour identifier rapidement les sources sonores ou formes de premier niveau, il reste à concevoir un mode de traitement aussi souple et efficace pour interpréter les formes de second niveau dont la structure spectrotemporelle est indépendante de la matière sonore.

Dessin 1 Dessin 2 Dessin 3 Dessin 4

Reconnaissance de dessins « sonorisés » à l'icophone.

Consigne : écoutez les quatre sons ci-dessous en vous efforçant de trouver le dessin qui correspond à chacun d'eux (réponse page 205, légende du son **4.33**).

Son 4.33b (4'') **Son 4.33c (4'')** **Son 4.33d (4'')** **Son 4.33e (3'')**

1. L'« icophone » est un synthétiseur de parole construit au LAM sur une idée de Leipp (voir Leipp, 1968, 1969).
2. Pour beaucoup d'auditeurs, la perception est immédiate, quasi inconsciente, et pourtant totalement imprévisible dans le contexte proposé.

3.10. Les ambiguïtés d'écoute et les illusions

> *Les illusions, ambiguïtés et autres anomalies visuelles qui ont suscité l'intérêt d'artistes et psychologues de la perception ne sont nullement des manifestations de lubies, pépins ou erreurs de conception arbitraires du système visuel humain. Elles résultent plutôt de l'application de puissants principes de déduction automatisés qui sont bien adaptés aux propriétés générales du monde naturel. Si nous sommes actuellement en vie, c'est parce que ces principes ont servi efficacement nos ancêtres.*
>
> Shepard, R. N., 1992, p. 218.

Illusions sonores

On connaît les figures ambiguës se prêtant à deux interprétations différentes, comme la tête de Boring[40], que l'on peut voir comme une jeune fille ou une vieille femme, ou encore le tableau de René Magritte, *Les Promenades d'Euclide* (1955), dans lequel une tour peut aussi être vue comme une rue en perspective selon l'interprétation que l'on donne au contexte. Dans tous ces exemples, c'est bien le sujet « regardeur » qui donne sens à la figure en décidant d'une interprétation. Nous proposons d'écouter un exemple similaire emprunté à l'art vocal qui joue sur l'ambiguïté perceptive entre le vibrato et le trille. En effet, le même ornement (Son **4.34**) peut être interprété tantôt comme un vibrato (Son **4.35a**), tantôt comme un trille (Son **4.35b**). Pour une présentation plus détaillée reportez-vous au chapitre 6, page 272.

Son 4.34 (3'')

Son ambigu

Son 4.35a (4'')

Le son 4.34 dans un premier contexte mélodique

Son 4.35b (5'')

Le son 4.34 dans un autre contexte

Les illusions fascinent par le décalage qu'elles opèrent entre les stimuli et leur interprétation. Confronté à des sons paradoxaux, l'auditeur ne peut rester écartelé entre différentes interprétations : il opte pour celle qui est la plus probable dans un contexte donné ou pour celle qui lui convient à un moment donné. L'acousticien qui connaît la structure acoustique du signal conclut à une illusion des sens au seul motif que le monde du signal et de l'analyse physique serait, selon lui, plus « vrai » que le monde des sens. Pourtant, en réponse au physicien Ohm qui considère comme une illusion le fait de fusionner en un son unique deux sons simultanés en rapport d'octave, Helmholtz (1874, p. 90) écrit : « l'oreille doit être seul juge en matière de sensation auditive, et on ne peut considérer comme illusion la manière dont elle perçoit les sons. » Plus récemment, Berthoz (1997, chapitre 13) affirme : « les illusions perceptives sont en réalité des solutions trouvées par le cerveau lorsque les informations sensorielles sont ambiguës, ou contradictoires entre elles ou avec les hypothèses internes qu'il peut faire sur le monde extérieur. On tient généralement l'illusion pour une erreur des sens (...), l'illusion (est) la meilleure hypothèse possible. » Nous présentons plusieurs exemples d'illusions au cours de l'ouvrage : illusions de perception de la hauteur de J.-C. Risset (chapitre 6, § 2.6 et § 3.5.2) ; illusion d'une voix virtuelle dans un chant polyphonique sarde (chapitre 9, § 3.5) ; illusion d'un sifflement se superposant à la voix dans la technique du chant diphonique « sygyt » (chapitre 6, § 2.7). Le lecteur pourra aussi lire avec profit la thèse de François-Xavier Féron : *Des illusions auditives aux singularités du son et de la perception,* 2006.

3.11. La transformation d'une forme dans une autre : le *morphing*

Le jeu est bien connu : comment passer insensiblement du beau visage d'Apollon à une tête de grenouille ? On procède par une série d'anamorphoses qui, à chaque étape, proposent une modification acceptable conduisant insensiblement vers la cible. La transformation est remarquablement accomplie par Grandville en sept étapes (figure 4.36). Pourtant, à bien y regarder, un trouble naît lorsque nous cherchons à repérer la transition critique entre les deux états : humain et batracien. En particulier le cinquième dessin (figure 4.37) extrêmement troublant, est celui d'un être sans nom : il n'existe pas d'être mi-humain mi-batracien.

40. La tête est reproduite sur le site anglais de Wikipedia, à l'entrée « Erwin Boring ». On la trouvera également, ainsi qu'un grand nombre d'autres exemples sur le site très bien documenté : « Hall_of_illusions » de l'université technique de Dresde.

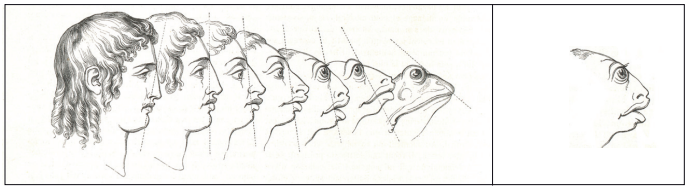


Figure 4.36 Transformation d'une forme dans une autre.

Grandville, J.-J., 1844, *Le Magasin pittoresque.* [Bnf]

Figure 4.37 Chimère…

Depuis la numérisation des sons, il est devenu aisé de réaliser des transformations équivalentes appelées *morphing*. Partant d'un son A on prélève des portions d'un son B qui remplacent les portions équivalentes du son A.

Au bout d'un certain temps, on arrive au son B. À l'écoute de la séquence **4.36** réalisée par Godøy[41], un trouble similaire à celui de la chimère grenouille-Apollon s'instaure dans les étapes centrales, lorsque l'instrument A, qui perd progressivement ses caractéristiques, n'est pas encore l'instrument B. La transformation acoustique est continue, mais il n'existe en réalité que deux formes qui peuvent faire sens : un Apollon ou une grenouille, un piano ou une contrebasse.

Son 4.36 (26″)

Morphing 1 (R. Godøy)

La transformation effectuée dans le Son **4.37**, très simple dans son procédé, est tout à fait saisissante. Au début, nous entendons une voix chantée articulant une syllabe voisine de « ala », dont la hauteur fondamentale chute très rapidement, qui disparaît pour faire place à un son de percussion musical ressemblant à une clochette.

Son 4.37 (10″)

Morphing 2 (X. Rodet)

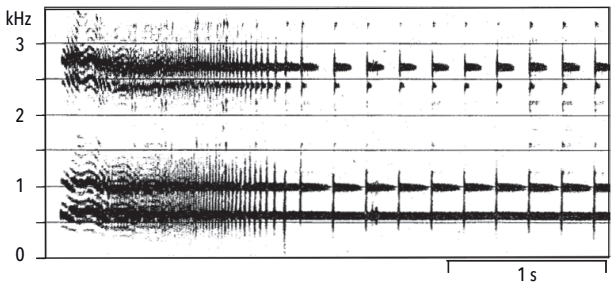


Figure 4.38 Analyse spectrale de la fin du Son 4.37.

Les auditeurs sont incrédules lorsqu'ils apprennent que cette métamorphose perceptive est le seul produit du ralentissement graduel des impulsions sonores du son vocal, ce que l'on voit clairement sur l'analyse de la figure 4.38. Nous savons que des impulsions de cadence rapide fusionnent perceptivement pour donner une « note » et qu'en deça d'un certain seuil temporel elles produisent une sensation discontinue (voir chapitre 3, page 118). Mais ce qui nous intéresse ici, c'est le fait que nous entendons tout d'abord une voix humaine dont la hauteur fondamentale descend, et que brusquement nous passons à un joli tintement de clochette. Une fois de plus, nous constatons que l'écoute, qui a pour finalité première de donner sens aux sons, identifie prioritairement la source acoustique. La rapidité d'appréhension dépend de la prégnance des formes en présence mais surtout des capacités de l'auditeur à anticiper, saisir et interpréter les données acoustiques de façon cohérente. Comme le montrent les recherches en psychologie cognitive, le cadre dans lequel nous effectuons l'interprétation de ces données est celui de la catégorisation.

41. Godøy, R., 1993, *Formalisation and epistemology,* thèse ; Det historiskfilosofiske fakultet, Oslo.

4. La catégorisation perceptive des sons et des séquences

4.1. La notion de catégorie

> *Without the ability to categorize, we could not function at all, either in the physical world or in our social and intellectual lives.*
>
> Lakoff, G., 1987, p. 6[42].

Selon le *Petit Robert*, catégoriser, c'est « classer, c'est-à-dire grouper, rassembler, mais également délimiter, séparer une collection d'objets ou de concepts ». Cette opération fondamentale de mise en ordre, d'organisation d'un grand nombre de données préoccupe aussi bien les philosophes que les naturalistes. Il existe donc différentes conceptions de la catégorisation et différentes manières de procéder, selon la nature des données à catégoriser et selon les finalités de l'opération.

S'il s'agit de catégoriser les concepts comme le fit Aristote[43], l'ensemble de l'édifice – nombre de catégories, définitions et exemplaires donnés en illustration – dépend étroitement du développement des idées à un moment donné. C'est dire que de telles catégorisations sont régulièrement remises en cause et alimentent le débat philosophique. S'agissant des éléments naturels, les systèmes de catégorisation développés au XVIIe et XVIIIe siècles sont plutôt de type taxinomique. Ce sont des édifices hiérarchiques rigoureux, fondés sur l'élaboration de critères objectifs mesurables. Tous les éléments d'un niveau donné sont inclus dans une catégorie de niveau supérieur. L'histoire des sciences montre toutefois qu'ils sont également remis en cause lorsque les connaissances évoluent et qu'il faut prendre en compte des données nouvelles.

Avec le développement des sciences cognitives, les théories de la catégorisation ont connu un renouveau considérable au XXe siècle et se sont illustrées dans le cadre de deux courants principaux.

Pour les uns (Jean Piaget, 1959), l'acquisition des connaissances se ferait dans le cadre d'une catégorisation apprise, logiquement construite, quasiment universelle, que les enfants doivent s'approprier au fur et à mesure de leur développement.

Pour d'autres (Eleanor Rosch, 1976 ; Danièle Dubois, 1993), la mise en place et la structure des catégories des connaissances serait une construction de la personne et s'établirait dans une interaction constante avec les données de la perception, donc selon des modalités dépendantes de l'expérience et des motivations des individus.

Selon d'autres encore (Jacqueline Bideaud et Olivier Houdé, 1989), les deux modalités de catégorisation coexisteraient et se développeraient concurremment chez l'enfant. Il est à noter toutefois que l'acquisition du langage fournit un outil puissant de catégorisation logique, et que, simultanément, l'analyse des verbalisations fait état d'une certaine diversité dans les modalités de catégorisation.

42. Sans la possibilité de catégoriser, nous ne pourrions pas fonctionner du tout, que ce soit dans le monde physique ou dans nos vies sociale et intellectuelle. (Traduction : M. C.)
43. L'essence, la qualité, la quantité, la relation, l'action, la passion, le lieu, le temps, la situation, la manière d'être. A. Cuvillier, *Vocabulaire philosophique*, Armand Colin, 1956.

Catégoriser

Un exemple de vie courante

Prenons un exemple familier : ranger un grand nombre d'objets après un déménagement. On fera vraisemblablement des tas différents avec les vêtements, le matériel de cuisine, les livres, les objets de toilette, c'est-à-dire selon l'usage que nous avons de ces objets. Si maintenant nous reprenons le tas des livres, nous pourrons hésiter entre un classement logique par matières ou par auteurs ou encore par dimensions si les étagères nous y contraignent ! On voit que les critères sont définis par celui qui catégorise et peuvent être remis en cause lorsque le besoin s'en fait sentir. Les critères sont « bons » lorsqu'ils permettent d'aboutir à un ordonnancement satisfaisant tout en offrant la possibilité d'accéder rapidement et sans hésitation à un élément particulier.

Catégorisation perceptive et langue

Sans le savoir nous pratiquons au quotidien une catégorisation hiérarchique. La langue en est le révélateur et aussi, en quelque sorte, l'élément structurant. Prenons un exemple commun : la « pomme ». Cette même pomme devient « un fruit » si à la fin d'un repas on nous demande de choisir entre « un fruit ou une pâtisserie », ou encore une « reinette » à l'étal d'un épicier qui a aussi des « goldens » et des « calvilles ». Ces trois termes d'usage courant – pomme, fruit, reinette – témoignent d'une organisation logique des connaissances : la catégorie « fruit » comprenant différents représentants, dont les pommes, lesquelles se déclinent selon plusieurs variétés comme la reinette.

Ces catégories paraissent évidentes, « données » avec les connaissances, voire quasi universelles. Un tel point de vue a été brutalement remis en question dans la deuxième moitié du XXe siècle, à la suite de travaux d'ethnolinguistes. En effet, l'analyse des langues de différentes cultures a mis à jour d'importantes disparités dans l'organisation des éléments naturels de l'environnement, des animaux, des plantes. L'exemple le plus connu est celui des Nenets[1], un peuple esquimau qui utilise une quarantaine de termes pour décrire l'état de la neige. Une telle abondance témoigne de l'importance qu'a cet élément dans la vie quotidienne de ce peuple, puisque la chasse et la pêche en dépendent. Elle révèle aussi que les Nenets structurent le monde sensoriel très différemment de nous.

De tels exemples ont conduit les chercheurs à s'interroger sur les modalités de constitution des catégories dites « naturelles », c'est-à-dire celles qui concernent les « objets » du monde environnant.

1. A. Kondratov, *Sons et signes*, p. 104, Mir, 1968.

4.2. La catégorisation prototypique : similarités, niveau de base et typicalité

Why do human culture categorize the world in the way they do ?[44]

Rosch, E. & Lloyd, B. (éd.), 1978, p. 1.

Eleanor Rosch et ses collaborateurs ont effectué plusieurs expériences pratiques de catégorisation portant sur des mots ou des dessins représentant des objets du monde réel (carotte, voiture, tulipe), afin d'étudier les relations entre ces objets et les catégories linguistiques (légumes, véhicules, fleurs). Les résultats des tests ont mis en évidence d'importantes différences entre la constitution de ces catégories et celles qui sont fondées sur des principes logiques comme les taxonomies des sciences de la nature. D'une façon générale une catégorie rassemble des objets « similaires » (en anglais : *family resemblance*), dont on peut quantifier le taux

44. Pourquoi les cultures humaines catégorisent-elles le monde de la manière dont elles le font ? Traduction : M. C.

d'attributs communs aux membres d'une même catégorie et celui des attributs contrastant deux catégories différentes. À partir de ces résultats ont été définies deux notions importantes : le « niveau de base » et la « typicalité ».

Organisation des catégories : le niveau de base

L'agrégation par similarité s'opère prioritairement au niveau de base. C'est celui des items les plus courants, les mieux partagés par les membres d'une communauté et ceux qui correspondent aux désignations en langue les plus courantes. Par exemple une pomme, un chat, une bouteille, une casserole sont des catégories d'objets constitués au niveau de base, pour lesquels un enfant acquiert rapidement le vocabulaire. Ce n'est que beaucoup plus tard qu'il apprendra à regrouper bouteille et casserole dans la catégorie « récipient », située au niveau supérieur, dit sur-ordonné. Par ailleurs, la distinction entre les catégories « bouteille de bordeaux » et « bouteille de bourgogne », qui relève d'un niveau plus expert, dit sub-ordonné, ne s'acquerra vraisemblablement que si le besoin s'en fait sentir. Le schéma de la figure 4.39 donne un exemple possible des trois principaux niveaux de catégorisation sonore.

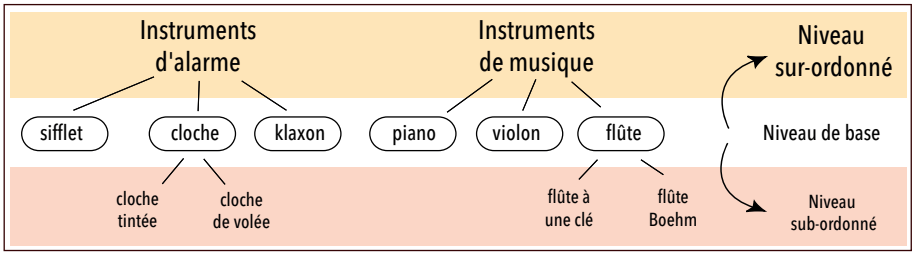


Figure 4.39 Exemples d'une organisation hiérarchisée de catégories sonores. Les niveaux « sur-ordonné » et « sous-ordonné » s'organisent de part et d'autre du niveau de base, niveau partagé par le plus grand nombre d'auditeurs.

Structure interne des catégories : le prototype

La deuxième notion importante de la catégorisation prototypique concerne la structure interne des catégories. Les travaux de Rosch ont montré aussi que les divers exemplaires d'une catégorie ne sont pas équivalents : certains sont de meilleurs représentants de la catégorie que d'autres. Voici une expérience simple. On demande à un groupe de personnes d'écrire très rapidement une liste de dix noms d'instruments de musique. En comptant les résultats, on voit que l'instrument le plus fréquemment cité est le piano. Ensuite viennent le violon, la guitare, la flûte, la clarinette, etc. Pour ce groupe, le piano est le meilleur représentant de la catégorie instruments de musique : il en est le « prototype ». Les autres instruments se situent à des distances variables selon leur degré de typicalité. Il est bien évident que cette organisation (figure 4.40) n'est valable que pour le groupe interrogé[45].

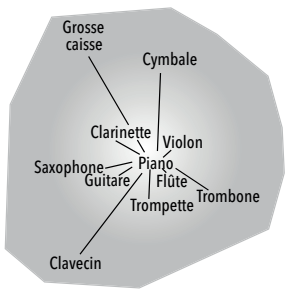


Figure 4.40 Exemple de structure de la catégorie « instruments de musique ». Le prototype est au centre. La position plus ou moins distante des autres instruments figure leur degré de typicalité tel qu'il est issu de l'expérience décrite.

45. En l'occurrence deux groupes d'étudiants en acoustique pour lesquels les résultats ont été très similaires : l'un en Master-Atiam, l'autre dans la classe d'acoustique du Conservatoire de musique de Paris. Nous répétons cette expérience chaque année avec des résultats très stables.

L'ensemble des données culturelles (les connaissances, les goûts et la pratique musicale) conditionnent à la fois la richesse et la diversité des éléments d'une catégorie donnée, donc sa structure.

Les deux modes d'organisation des catégories, typicalité horizontale et structuration verticale de part et d'autre du niveau de base, répondent à des principes d'économie cognitive pour le codage et la reconnaissance des éléments des catégories. On voit que l'organisation des données n'est pas soumise à une logique d'abstraction par inclusions successives, mais qu'elle est plutôt structurée en vue d'une efficacité perceptive optimale. Au niveau de base se trouvent les classes des éléments les plus pertinents pour un individu, incluant des prototypes qui ont pour particularité d'être à la fois les plus représentatifs de leur catégorie et les plus aisément discriminables du point de vue perceptif.

Danièle Dubois et ses collaborateurs (1993, 1997) ont repris et développé la catégorisation prototypique pour explorer des perceptions sensorielles complexes dont l'expression en langue est soit défective (odorat), soit hétérogène et fort dépendante des sujets (ouïe).

La pratique expérimentale de la catégorisation libre que nous allons présenter a été largement développée au Laboratoire d'acoustique musicale. Elle s'est révélée une méthode de choix pour étudier la perception des sons complexes, tant du point de vue de leur identification que de leur qualification.

4.3. La catégorisation et la perception sonore

> *There is an unmistakable difference between « seeing » that two things look similar and « judging » that they belong in the same category.*
>
> Neisser, U., 1967, p. 95[46].

Ulric Neisser envisage deux processus de catégorisation : l'un dit perceptif par lequel sont traitées les similitudes entre objets, l'autre dit conceptuel par lequel le sujet décide de l'appartenance à une catégorie par une opération constructive[47].

Quelles sont les parts respectives des critères acoustiques et des critères culturels dans notre catégorisation du monde sonore ? Comment interfèrent les données perceptives et les mots que nous utilisons pour décrire les phénomènes sonores ? L'expérience de catégorisation libre présentée ci-dessous montre l'existence de regroupements au plus près de la typologie acoustique.

4.3.1. Deux expériences de catégorisation libre avec des sons brefs, hors contexte

Dans un test de catégorisation libre, chaque individu décide du nombre et de la composition des catégories. L'analyse mathématique des résultats fournis par un groupe de sujets donne de précieuses indications sur la pertinence des catégories formées ; l'analyse des verbalisations en éclaire l'interprétation. D'une part, le vocabulaire employé fournit des données précieuses pour orienter l'étude acoustique et, d'autre part, la formulation verbale éclaire l'expérimentateur sur l'implication personnelle des auditeurs lors de l'élaboration de leurs perceptions (voir aussi chapitre 7, § 4.3).

46. Il y a une indubitable différence entre se rendre compte que deux choses ont l'air semblables et juger qu'elles appartiennent à la même catégorie. (Traduction : G. Bloch.)
47. « One does not simply examine the input and make a decision ; one builds an appropriate visuel object ». Neisser, *op. cit.* ; p. 94. On ne peut se contenter d'examiner ce qui nous arrive et prendre une décision : on doit construire un objet visuel adéquat. (Traduction : G. Bloch.)

L'expérience dont nous rendons compte a porté sur les deux collections de sons A et B (29 sons chacune), déjà analysées au § 3.3. Chaque son est une icone sonore affichée sur l'écran de l'ordinateur. L'auditeur écoute tout d'abord la totalité des sons, puis il reçoit comme consigne de « faire des groupes de sons similaires ». Il peut réécouter les sons autant de fois qu'il le désire et déplacer les icones pour former le nombre de groupes qui lui convient. À l'issue de cette opération, l'auditeur doit expliquer les particularités ou les motivations qui ont présidé à la constitution de chaque catégorie. L'analyse et le dépouillement des résultats sont présentés dans l'encadré ci-dessous.

Expérience de catégorisation libre : dépouillement des résultats

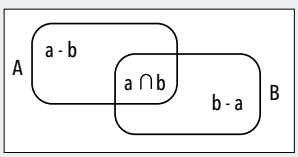


Figure 4.41 Expression de la similarité (s) entre deux objets A et B.

Catégoriser, c'est comparer : évaluer les ressemblances et les dissemblances entre des items, apprécier un contraste. Soient par exemple deux d'objets complexes A et B, et les attributs a et b qui caractérisent chacun d'eux. L'opération qui permet d'estimer la similarité entre les deux objets doit prendre en compte trois groupes d'arguments :

1/ les attributs communs à A et à B ;

2/ les attributs que A possède et que B ne possède pas ;

3/ les attributs que B possède et que A ne possède pas.

Selon cette théorie (Barthélémy et Guénoche, 1988), la similarité (s) est calculée comme une combinaison linéaire de leurs caractères communs et de leurs différences (figure 4.41).

L'algorithme est d'abord appliqué au groupe des sujets pour isoler ceux qui ont des réponses trop divergentes, puis à chacun des éléments du corpus. On peut représenter les résultats de différentes façons.

Dans la représentation arborée de la figure 4.42 page suivante les feuilles terminales numérotées – ici les sons – forment des bouquets plus ou moins denses. La donnée pertinente pour l'interprétation d'un arbre est la longueur des segments, qu'il s'agisse des stimuli ou des catégories.

Les sons 21 et 26, très proches, forment avec le son 7 la catégorie « voix humaine » à laquelle s'agrège tardivement le n° 28, voix chantée (nœud N).

De la même façon, cinq sons forment la catégorie des « cris d'animaux », laquelle rejoint le groupe « voix humaines ». Toutefois la longueur du segment de rattachement (passant par P) indique que ces deux catégories sont bien distinctes.

Expérience de catégorisation libre des sons de la collection A. Six catégories principales ont été formées (figure 4.42). Les sons de « voix humaine » et ceux des « cris d'animaux » constituent deux catégories distinctes mais proches : il s'agit de sons produits par des êtres vivants.

Les deux catégories « sons percussifs », et « instruments de musique » rassemblent chacune des sons similaires du point de vue acoustique.

Enfin, les « signaux d'avertissement » regroupent des sons sur une base fonctionnelle et « machines » sur celle d'un type de production (moteurs, percussion mécanique). C'est donc une catégorisation mixte, « construite » par les sujets, combinant des critères acoustiques et des critères relatifs à l'aspect fonctionnel des sources. Ainsi, le son 23, pourtant très voisin du 12 (flûte) sur le plan acoustique, se retrouve dans les signaux d'avertissement, car il a été reconnu comme signal de train. Par ailleurs, la cloche (son 5) oscille selon les sujets entre les catégories « signaux » et « instruments de musique » ; la voix chantée (son 28) oscille entre « voix humaine » et « instruments de musique ».

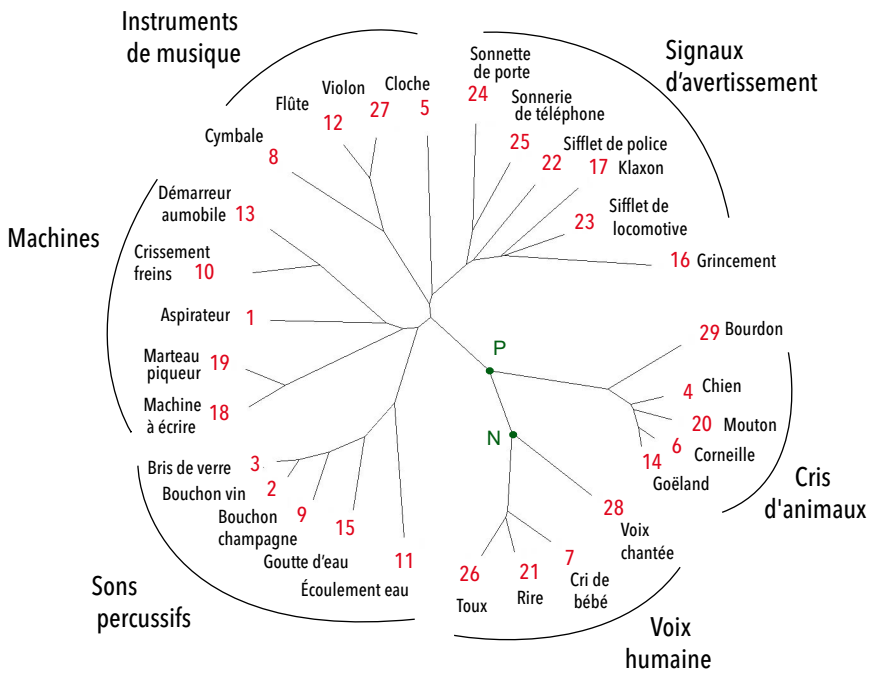


Figure 4.42 Représentation arborée de la catégorisation libre de 29 sons de l'environnement quotidien réalisée par 16 sujets d'âges et d'origines très variés, sans compétence particulière en acoustique ou en musique. (voir aussi la figure 4.8, page 156, pour l'analyse sonagraphique des sons)

Son 4.4 (22'')

Collection A, 29 sons

Expérience de catégorisation libre des sons de la collection B. La tâche de catégorisation libre de ce corpus s'est révélée beaucoup plus difficile à accomplir que la précédente, car les auditeurs ne peuvent pas désigner les sons par un nom de source. Ils s'efforcent alors d'identifier les actions mécaniques (grincement, craquement) ou les gestes (frottement, grattement, déchirement) qui pourraient être à l'origine du son (figure 4.43). Cette fois la catégorisation s'opère explicitement sur

des critères acoustiques, avec une prédominance nette des caractéristiques micro-temporelles car les sons sélectionnés sont brefs et dépourvus de raies spectrales. On remarque aussi l'emploi par les sujets de nombreux termes qualificatifs dans leur vocabulaire de description des catégories : sons doux, sourds, continus, ainsi que des termes hédoniques comme agréables, stridents, qui ont trait à l'effet du son sur l'auditeur (voir Dubois, 1993 ; Guyot, 1997).

Son 4.5 (21'')

Collection B, 29 sons

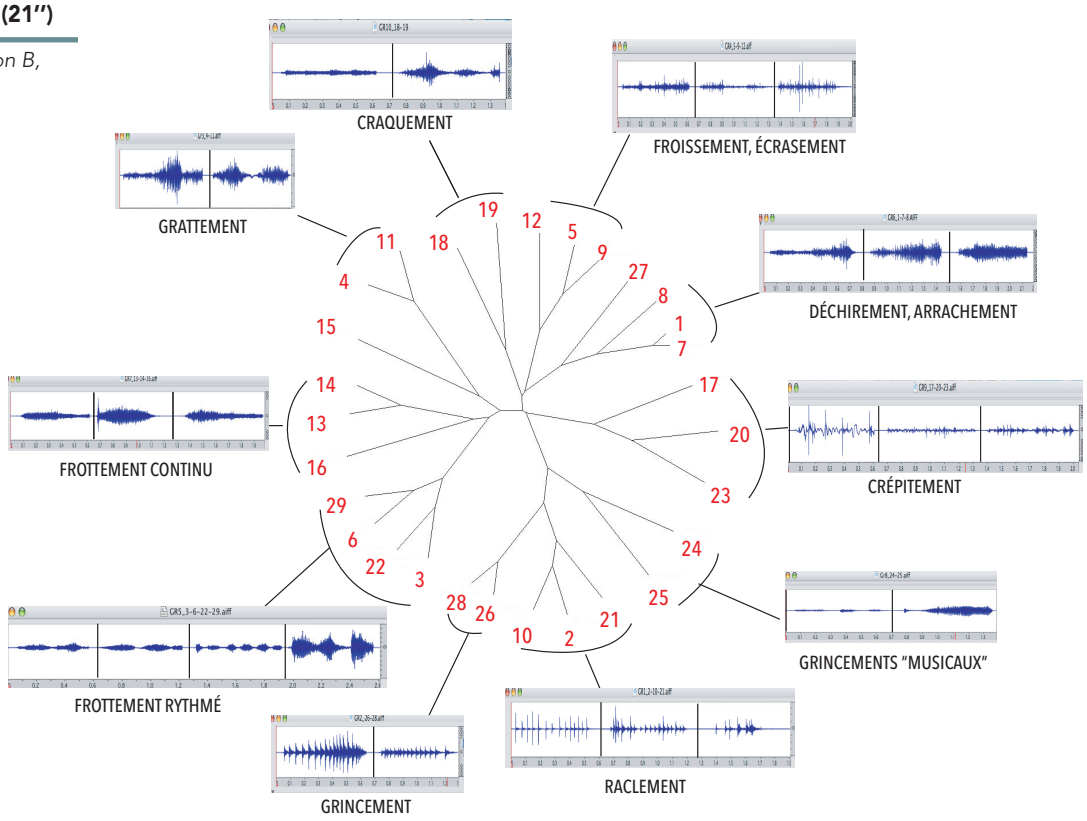


Figure 4.43 Catégorisation des sons de la collection B, avec le signal temporel des items. (Voir aussi la figure 4.9, page 156, pour l'analyse sonagraphique des sons)

Bilan des deux expériences de catégorisation. Étant donné que les deux expériences décrites se déroulent dans des conditions similaires, on peut penser que les différences observées dans la tâche de catégorisation sont imputables aux corpus sonores. Par ailleurs, le choix de stimuli brefs, hors contexte, porte à croire que l'écoute causale est prioritairement sollicitée dans les deux cas.

Pour la collection A, la tâche est réalisée très rapidement (moins de 15 min) et les catégories sont formées sur une base combinant la sémantique et l'acoustique. Pour les sons de la collection B qui ne peuvent pas être attribués à des sources connues, le temps passé à la catégorisation est beaucoup plus long (de l'ordre de 1 h). Les auditeurs s'efforcent d'analyser les modes de production et les termes employés pour la description des catégories révèlent une grande diversité de modalités perceptives, la majorité d'entre eux désignant des actions (frottement, grattement, déchirement) ou leurs effets (grincements)[48].

Dans les deux cas, les auditeurs s'efforcent d'attribuer une causalité au son – une *source* ou une *action* –- mais les deux types d'écoute sont très différents. Lors de l'expérience avec les sons de la collection A, l'écoute est de type indiciaire : l'auditeur stoppe la lecture du son dès qu'il reconnaît la source et qu'émerge le mot la désignant. À l'opposé, l'échec d'une reconnaissance immédiate qui se produit avec les sons B fait basculer l'auditeur dans un mode analytique qui mobilise toute son attention. Il reprend à plusieurs reprises la comparaison des sons entre eux, s'attache à les discriminer tout en s'efforçant d'élaborer un vocabulaire lui permettant de mémoriser les caractéristiques repérées. Certains sujets, ayant passé le test B en premier, ont eu du mal à commencer le test A tant le contraste entre les deux modes d'écoute est grand. Ajoutons que les auditeurs passionnés de musiques électroacoustiques, qui ont développé une écoute des qualités des sons indépendamment de leur étiquette sémantique, montrent plus d'aisance dans la réalisation du test B – l'un d'eux a même pris pour parti de catégoriser aussi les sons du test A sur des paramètres acoustiques. Il est vrai qu'une expérience en laboratoire, au cours de laquelle l'auditeur n'est pas soumis aux contraintes d'une communication sonore en temps réel, autorise toutes sortes de stratégies d'écoute.

4.3.2. Les comportements d'écoute

Les situations que nous vivons quotidiennement, pour lesquelles nous avons développé différents comportements d'écoute, sont bien éloignées de celle de l'expérience qui vient d'être présentée. Habituellement, tous les sens coopèrent dans la mémorisation de scénarios types, ceux sur la base desquels nous sommes capables d'anticiper la majorité des événements sensoriels qui peuvent se produire : odeurs, images, mouvements et sons. À vrai dire, il ne s'agit plus de *sons* mais de parole, du bruit d'un ventilateur, d'une porte qui claque, d'un poste de radio qui diffuse de la musique ou d'un merle qui siffle, c'est-à-dire de phénomènes sonores s'inscrivant dans une scène globale et cohérente qui fait sens. La catégorisation joue un rôle fondamental dans la façon dont chacun de nous construit les relations entre l'audition et les autres sens, ainsi que dans l'organisation des données acoustiques et syntaxiques des séquences sonores relatives aux deux situations majeures de communication entre humains que sont la conversation et l'écoute de la musique.

Nous possédons en mémoire une typologie perceptive des sources et des prototypes catégoriels associés. Et puisqu'à l'exception des sons qui nous alertent (voir encadré page suivante), nous sommes libres des choix de notre écoute, il existe de fait une très grande diversité dans les modalités de catégorisation de l'univers sonore, diversité que l'on peut tenter d'explorer expérimentalement à l'aide de tests d'écoute contrôlés.

4.3.3. Les tests d'écoute

Analyser l'écoute est une tâche difficile, car aucun indice visuel ne permet à l'expérimentateur d'observer l'activité auditive d'un sujet. À la seule interpellation « Écoutez », nous croyons tous entendre la même chose, or il n'en est rien. C'est seulement par l'interrogation minutieuse de chaque auditeur que l'on peut mettre en évidence la diversité des écoutes, tant dans leur contenu que dans leur stratégie. Par ailleurs, l'interrogatoire, qui déjà ne peut intervenir que sur la trace mémorisée du son, se heurte aussitôt à un obstacle de taille : l'absence d'un vocabulaire défini et consensuel pour exprimer verbalement ce qu'on a entendu. Seuls les experts qui ont besoin de communiquer leur expérience auditive – professeurs d'instruments, techniciens du son, acousticiens – développent une terminologie et des expressions

48. Les « étiquettes » des catégories de la figure 4.43 sont données par l'expérimentateur, sur la base d'une confrontation entre le signal acoustique et les termes utilisés le plus fréquemment par les auditeurs du test.

Sons subis ou sons choisis

Lorsqu'un son sollicite notre oreille, il déclenche – que nous l'ayons ou non désiré – une tentative d'identification qui s'effectue à notre insu, en cohérence avec toutes les autres données perceptives. Parmi les sons de l'environnement, seuls les sons surprenants par leur intensité ou par leur incongruité éveillent notre attention : dans tous les cas, nous les subissons.

Pourtant, nous pouvons aussi choisir d'écouter intentionnellement certains sons plutôt que d'autres. Prenons l'exemple banal d'un café animé à l'heure du déjeuner. Il s'y trouve des bruits de vaisselle, de machine à café, la musique de fond, des conversations, les annonces des garçons : toutes sortes de formes sonores auxquelles la plupart des personnes présentes ne prêtent pas attention. Imaginons que vous êtes là pour enregistrer le paysage sonore du lieu. Vous fermez les yeux et vous prenez conscience de la richesse sonore de l'espace dans lequel votre oreille voyage avec plaisir. Soudain, une personne que vous connaissez entre et engage une conversation avec vous. Le théâtre sonore ambiant disparaît subitement au profit de la voix de votre interlocuteur sur laquelle votre écoute est maintenant toute entière focalisée. Du moins c'est ce qu'il peut croire, car il ne saura jamais comment vous l'écoutez, si vous écoutez ce qu'il dit (sémantique) ou comment il le dit (qualitatif) puisque la voix en tant que source et la parole qu'elle transmet peuvent donner lieu à une grande diversité de modes d'écoute choisis.

Nous savons que les aspects sémantiques et les aspects qualitatifs sont antagonistes (voir § 2.2). Ils se déroulent à des échelles de temps différentes et ne portent pas sur les mêmes éléments du signal. Dans les sons de l'environnement, il y a conjonction entre le matériau sonore, le signal et l'information portée par sa forme acoustique. Les deux font bloc, au point qu'il est extrêmement difficile d'occulter l'événementiel pour accéder à l'hédonique : il faut le vouloir, il faut s'y entraîner. En parole, ces deux aspects sont dissociables. Les caractéristiques intrinsèques du matériau sonore rendent compte de la source et sa modulation porte les formes sémantiques de la parole. Par nécessité et par habitude, le décryptage du contenu sémantique du signal l'emporte le plus souvent sur les aspects qualitatifs de la voix auxquels pourtant les auditeurs sont sensibles, mais au second plan.

Contrairement aux sons de l'environnement et à la parole, l'écoute de la musique se pratique rarement de façon inopinée car, généralement, la situation et le lieu préparent aux activités d'écoute musicale. Toutefois, l'extrême diversité des créations contemporaines se heurte souvent à l'incompréhension des auditeurs en raison précisément des habitudes d'écoute développées au contact de la grande masse de musique tonale mélodico-harmonique. On connaît l'injonction de Pierre Schaeffer « travailler son oreille », et s'entraîner à l'écoute, dite « réduite », c'est-à-dire l'écoute du son pour lui-même.

verbales spécialisées témoignant d'une écoute raffinée et précise des qualités des sons, qui sont spécifiques à chaque domaine.

Tout au long de cet ouvrage, nous usons abondamment d'exemples sonores offrant au lecteur la possibilité d'analyser son écoute personnelle afin de confronter ses impressions avec les interprétations que nous proposons. L'étape suivante consiste à organiser un test avec plusieurs sujets pour obtenir des résultats généralisables à un groupe d'auditeurs, ce qui nécessite d'établir un protocole contrôlé et reproductible. Or, l'élaboration d'un test implique d'effectuer des choix qui vont nécessairement orienter l'écoute des auditeurs : choix des types de stimuli (synthétiques ou naturels), de leur durée et de leur nombre ; choix du type de tâche à effectuer (comparaison par paires, échelles sémantiques, catégorisation libre, classement ordonné)[49] ; choix des conditions d'écoute (casque, enceintes) ; choix de la consigne verbale. Pour l'exploration d'une nouvelle expertise d'écoute, il est

49. Pour un développement voir Bonnet, C. (1986). *Manuel pratique de psychophysique.* Paris : Armand Colin (254 p.).

prudent de commencer par un test peu directif, comme la catégorisation libre qui offre à l'auditeur la possibilité de mettre en œuvre indifféremment une stratégie holistique (globale) ou une stratégie paramétrique (dimensionnelle).

4.4. La catégorisation libre associée aux verbalisations

La catégorisation libre associée à des verbalisations spontanées se révèle d'un grand intérêt pour explorer la perception de la qualité sonore de scènes sonores ou d'extraits musicaux, à condition de s'assurer la collaboration de linguistes intervenant à toutes les étapes de l'expérience. En premier lieu pour définir la consigne verbale du test, ensuite pour établir le recueil du corpus verbal complet, enfin pour l'analyse et l'interprétation de celui-ci. C'est un travail long et rigoureux qui se révèle d'un grand intérêt pour l'interprétation des résultats. Il faut dire que bon nombre d'expérimentateurs s'autorisent à interpréter eux-mêmes les verbalisations, en se contentant de lister les adjectifs qualificatifs sans tenir compte du contexte et en posant implicitement que les termes utilisés par les auditeurs sont sans ambiguïté.

Nous avons conduit plusieurs expériences au LAM, en collaboration avec Danièle Dubois, pour évaluer la qualité sonore d'instruments de musique – guitare, clavecin, violons, archets de violon – ou pour étudier la perception de scènes sonores d'environnement urbain : tantôt pour estimer la « gêne », tantôt pour comparer des systèmes d'enregistrement sonore (voir chapitre 7, § 4.4.2). Selon les tests, différentes précautions s'imposent. Pour les séquences musicales, les instruments doivent être accordés au même diapason, joués par le même musicien et enregistrés strictement dans les mêmes conditions. Dans le cas de scènes sonores de l'environnement (Vogel, 1999), il est impératif d'effectuer des prises de son simultanées, car la première écoute est toujours une écoute de découverte causale du contenu de la séquence[50]. L'exploration des qualités, l'écoute des différences spécifiques ne peut valablement s'appliquer qu'à des séquences dont le contenu événementiel est identique. Les catégorisations de scènes urbaines ont montré en particulier que les séquences comportant des événements sonores identifiables étaient systématiquement regroupées et souvent jugées plus agréables que celles dont le bruit de fond, de niveau sonore équivalent, était « amorphe » (Maffiolo, 1999). Les verbalisations des scènes dites événementielles incluent des noms de sources – autobus, mobylette, klaxon, bruits de pas etc. – alors que pour les séquences dites amorphes, les auditeurs ont recours à des termes descriptifs comme « sons continus », « bruit de basse fréquence », « grondement ». L'analyse linguistique du corpus verbal a permis de préciser les différences de constructions perceptives des sujets dans les deux situations et a révélé une grande richesse dans la structure des catégories ainsi produites, catégories qui portent non seulement sur des propriétés du signal, mais également sur des activités, des marques du temps, ce qui peut permettre d'expliquer des comportements d'intolérance à des séquences de bruits dont la mesure d'intensité physique n'est pas élevée au regard des normes admises (Dubois et coll., 2005).

4.5. La catégorisation des sons de la langue et de la musique

Les sons musicaux et les sons de la langue donnent lieu à des catégorisations d'une grande richesse. Au cours de notre existence, nous entendons une quantité innombrable de voix humaines sur la base desquelles nous constituons des catégories selon l'âge et le sexe : femmes, hommes, enfants. D'autres catégories regroupent les sons propres à une langue donnée, avec toutes les variantes des accents locaux.

50. Nous avons constaté, lors d'une tâche de comparaison de techniques de prise de son, que les auditeurs devaient « épuiser » l'information événementielle des séquences avant d'être capables de basculer dans une évaluation qualitative.

Nous pouvons également imaginer des catégories phonologiques – les voyelles, les consonnes plosives, les chuintantes – à l'intérieur desquelles se retrouvent tous les types de voix. De façon analogue, les musiciens structurent leur monde sonore en constituant des catégories selon les instruments, la hauteur musicale ou la sonorité, c'est-à-dire en fonction des besoins propres à une pratique musicale donnée. La richesse et la pertinence des prototypes de chaque catégorie, qui dépendent de l'expérience personnelle, conditionnent à la fois la rapidité et la sûreté d'appréhension d'une musique, mais aussi le « goût » et les appréciations portées sur la qualité sonore des instruments. Selon Danièle Dubois (1993, p. 49) le prototype, qui rassemble le maximum de traits communs à tous les exemplaires d'une catégorie, n'a pas d'existence réelle. Pourtant, pour un individu donné, il est une référence mentale puissante et incontournable.

4.5.1. Du son à l'instrument de musique : catégories et prototypes

> *Ainsi le terme violon, dans l'indication « un sol de violon », n'est pas moins abstrait que la valeur désignée par le symbole « sol ». On a retenu, en oubliant le reste, ce qui pouvait être commun à tous les violons possibles.*
>
> Schaeffer, P., 1966, p. 317.

Ce qui, selon l'expression de Pierre Schaeffer, est « commun à tous les violons possibles », c'est-à-dire les traits partagés par tous les exemplaires sonores du violon, est le prototype identitaire de l'instrument. Il faut toutefois compléter la définition par l'ensemble des traits qui opposent les sons de violon à ceux des catégories adjacentes. Du point de vue acoustique, l'identité causale du violon est une combinaison entre le mode de production sonore (la corde frottée) et la réponse spectrale de la structure (la caisse, le chevalet, le cordier et le manche). La réalité est plus compliquée, car un instrument ne correspond pas à une signature acoustique unique. Ce n'est qu'au terme d'une longue pratique de la musique que l'on devient capable de reconnaître à l'oreille les instruments d'un orchestre tant leur variabilité est grande. Un instrument de musique est au service de l'expressivité musicale et, en tant que tel, doit se prêter à d'infinies variations sonores tout en restant reconnaissable. Pour approcher l'identité acoustique d'un instrument, il faut donc mimer ce que font les auditeurs en constituant un prototype qui représente en quelque sorte l'ensemble des corrélations entre les variations intrinsèques dues au jeu – intensité, durée, tessiture – et celles des réponses acoustiques de la structure. Aux caractéristiques acoustiques proprement dites il faut ajouter le style de jeu, le type de musique, les ornementations qui contribuent aussi pour une grande part à l'identité sonore de l'instrument. Enfin, pour un musicien, la catégorisation instrumentale inclut certainement les possibilités sonores mais peut-être plus encore la façon dont ces possibilités lui permettent d'interpréter la musique.

Chaque auditeur élabore ainsi ses catégories personnelles dont la structure se transforme au cours de son expérience d'écoute. Les prototypes sonores instrumentaux – piano, violon – qu'il a constitués lui servent de références pour la reconnaissance des sources et pour leur appréciation qualitative.

À titre d'exemple nous proposons, figure 4.44, deux représentations possibles de catégorisations comparées du piano et du clavecin, l'une vue par un acousticien (en haut) et l'autre par un musicien (en bas).

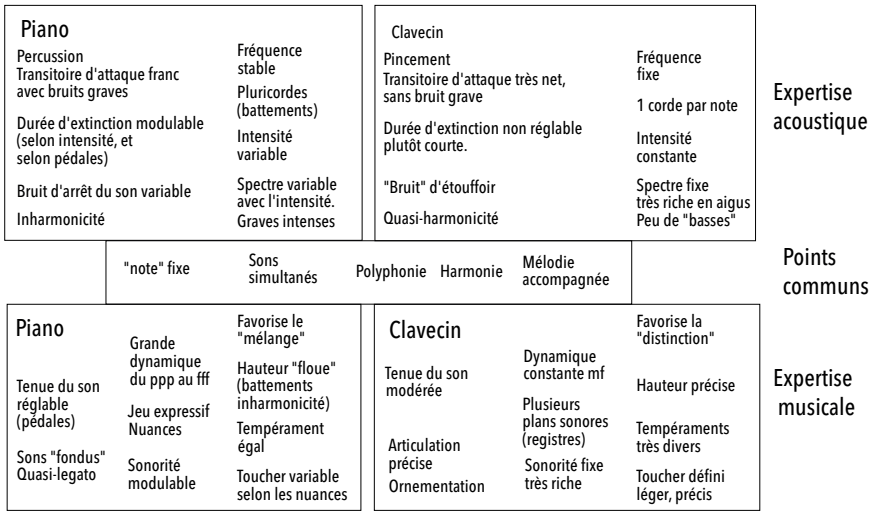


Figure 4.44 Catégorisation comparée des traits caractéristiques du piano et du clavecin, vue par un acousticien (en haut), et par un musicien (en bas).

Un acousticien construit une catégorisation qui s'attache aux caractéristiques de la production sonore, comme le mode d'excitation, la durée du transitoire et son contenu, l'harmonicité du spectre, alors qu'un musicien sera sensible de façon prioritaire aux possibilités que lui offre son instrument pour varier les sons dans le jeu legato ou staccato (donc au toucher) pour réaliser des nuances d'intensité, fondre les sons ou au contraire les articuler avec netteté.

Pour un type d'instrument donné, la catégorisation sonore d'un interprète combine aussi de façon indissociable son expérience physique du jeu, c'est-à-dire les actions qu'il effectue sur l'instrument, et les décalages qu'il constate entre la qualité sonore qu'il perçoit et celle qu'il s'attendait à produire.

4.5.2. La perception catégorielle : notes et syllabes

Contrairement à la catégorisation prototypique présentée au § 4.2, page 183, la perception catégorielle implique un système musical ou phonologique, extérieur aux individus qui doivent l'apprendre et s'y conformer. Pour un grand nombre de sons périodiques utilisés en musique, la sensation de hauteur est liée au paramètre de la fréquence[51]. La sensibilité auditive aux changements de fréquence est telle qu'il est possible de discriminer, dans les meilleures conditions, des intervalles de l'ordre du 1/300 d'octave (voir chapitre 3, page 104). Or, le système des hauteurs de la musique occidentale, qui ne comprend que douze sons à l'octave, nous conduit à développer une perception catégorielle des hauteurs musicales avec un échelon beaucoup plus grand que le minimum de discrimination sensorielle.

Prenons un exemple. Par convention, la note la_3 a pour fréquence 440 Hz. Augmentons légèrement la valeur de la fréquence : nous percevons la montée du son, toutefois la « note » proprement dite ne change pas : c'est un la_3 un peu haut. Si la fréquence continue d'augmenter, il arrive un moment où l'écart par rapport à la position convenue n'est plus acceptable : cette note devient un $la\#_3$. À la variation continue de la dimension de la fréquence correspond une perception musicale discontinue, par catégories de notes (écoutez l'expérience du Son **8.1** du chapitre 8).

51. Cette relation peut être mise en défaut avec certains sons de synthèse.

La catégorie des « *la* » renferme donc plusieurs sons de fréquences voisines dont les écarts par rapport à la note cible ne dépassent pas une valeur critique que l'on peut déterminer expérimentalement[52]. Il est clair qu'une telle catégorisation relève d'un apprentissage spécifique, propre à une culture musicale donnée, et qu'elle est renforcée par l'emploi de noms pour désigner les sons de l'échelle sonore. Cette pratique peut entraîner un conditionnement de la perception si poussé que certains musiciens perçoivent les notes de la gamme chromatique en toutes circonstances : en entendant des bruits complexes, des sons glissés, y compris pour des musiques dont les échelles sont très différentes de la nôtre. On montre de façon similaire que la discrimination des sons phonétiques d'une langue (voyelles et consonnes) est catégorielle.

Les frontières entre les items, d'une perception catégorielle sont bien définies, n'autorisant que peu de recouvrement, sous peine de générer des erreurs. Du point de vue de la richesse de tous les modes de perception possibles, c'est un processus réducteur qui « formate » la saisie perceptive des sons, bien qu'il soit possible d'y échapper en partie. En contrepartie, la perception catégorielle accélère le processus de reconnaissance des éléments appris en limitant le nombre des choix à effectuer. Elle permet en outre de constituer une collection d'items repérables se prêtant à une combinatoire de complexité variable, voire infinie. En musique comme en langue, la perception catégorielle donne seulement accès aux « briques » sonores de la construction du système. L'interprétation du sens, l'appréhension des structures et les jeux d'ambiguïtés se situent à d'autres niveaux d'élaboration.

4.5.3. Les ambiguïtés perceptives de la voix chantée : voyelles ou qualité vocale ?

La voix humaine tient une place singulière parmi les sources sonores, tant par son importance dans notre vie que par ses particularités acoustiques. Avec sa voix, qui est une source sonore (voir note 26), une personne peut faire varier toutes sortes de qualités : la hauteur, l'intensité, la durée, la sonorité (en donnant plus ou moins de souffle ou d'harmoniques et en changeant les voyelles). La voix transmet aussi la parole, le sens d'un discours. Selon le contexte (enseignement, concert, conversations à voix multiples), l'écoute de la voix peut être du type *reconnaissance* du message parlé, ou du type *qualification* de la voix du locuteur. Cette dualité est particulièrement forte dans l'exemple de la voix chantée pour laquelle les qualités du timbre vocal relèvent à la fois des voyelles – une voyelle est un timbre – et du contenu harmonique du son glottique (voir le chapitre 9).

Voyelle et chant

Son 4.38 (6'')

Son 4.39 (22'')

Voici une petite expérience illustrant des ambivalences possibles. Écoutons deux fragments très courts que nous avons prélevé dans une prestation chantée en prêtant attention aux voyelles (Son **4.38**) et en les notant par écrit. Ensuite écoutons l'exemple suivant (Son **4.39)** qui reprend les deux mêmes fragments, suivis des deux phrases musicales dont ils sont extraits. La surprise est grande d'entendre que la première phrase, qui commence sur le dernier son entendu, ne donne plus à entendre la même voyelle.

Il s'agit d'un bel exemple d'adaptation catégorielle au contexte. Le fait que les deux sons de l'exemple sonore **4.38** soient brefs favorise une discrimination fine, si bien que nous identifions les voyelles comme relevant de catégories distinctes[53]. À l'écoute de la phrase complète (Son **4.39**), l'interprétation du sens global du texte l'emporte sur la discrimination perceptive et la deuxième voyelle change de caté-

52. Voir E. Burns, « Intervals, scales, and tuning », in Deutsch, *The psychology of music*, 1999, p. 215-264.
53. Lisez le commentaire des sons page 205.

gorie. Mais ce qui est le plus remarquable est le fait que, ayant compris que le chanteur dit le même texte, nous pouvons, lors d'une troisième écoute, interpréter maintenant les différences acoustiques perçues en termes de qualités vocales. Nous entendons un chanteur qui interprète la première phrase avec une voix dite couverte et la deuxième avec une voix dite normale. Pendant l'expérience, nous avons basculé d'une perception catégorielle phonétique à celle d'une évaluation de qualité vocale.

4.6. Quelques réflexions sur la diversité des écoutes musicales

Les musiciens s'entraînent à l'écoute tout au long de leur vie. Dans l'enseignement musical, plusieurs exercices apprennent à se concentrer sur des éléments du vocabulaire (notes, rythmes, accords) qu'il faut parvenir à caractériser indépendamment des instruments, donc du matériau sonore. D'autres exercices, au contraire, entraînent à la discrimination des timbres, c'est-à-dire à écouter les caractéristiques intrinsèques au matériau sonore. La musique travaille donc indifféremment les écoutes de type analytique ou global, conceptuel ou sensible, et bien sûr esthétique. Très vite, des habitudes d'écoute s'installent et facilitent l'appréhension des œuvres d'un style donné, au risque de déformer ou d'interdire l'accès aux musiques d'autres cultures. Les exercices d'écoute pratiqués dans l'enseignement tendent majoritairement à renforcer des écoutes efficaces et spécialisées plutôt qu'à élargir et à diversifier les pratiques.

- L'écoute des *notes* qui est valorisée en Occident, est une écoute catégorielle favorisée et renforcée par la pratique de dictées. Entendre des notes focalise sur une écoute de type grammatical et phonétique, au détriment d'une perception plus globale (et peut-être plus sensible). Ce type d'écoute, qui peut développer une grande sensibilité à la justesse des intervalles d'un système donné, tend à occulter ou à interpréter comme faux, déviant, tout ce qui s'en écarte.
- L'écoute *musicologique* se développe avec la connaissance d'un large répertoire d'œuvres. Chaque musicien se constitue des grilles d'écoute grâce auxquelles il sait, où et comment, saisir dans la continuité sonore les indices qui lui permettront d'appréhender rapidement l'articulation des niveaux structurels d'une œuvre et ses caractéristiques stylistiques.
- L'écoute de *l'interprétation* implique que l'œuvre jouée fasse partie du répertoire de l'auditeur qui est alors capable d'apprécier les subtils écarts de justesse, de phrasé et de rythme par lesquels s'exprime un interprète, par rapport à une norme écrite ou par rapport à la tradition orale.
- L'écoute *polyphonique*, d'une grande complexité, est rarement d'un abord spontané et requiert aujourd'hui un effort constant pour se développer dans un environnement culturel où le genre de la mélodie harmonisée est largement dominant. Le suivi auditif des voix simultanées d'une polyphonie est facilité par le repérage d'indices (séparés ou concomitants) qui sont la tessiture, le timbre, les décalages temporels, la directivité spatiale et la logique mélodique.
- L'écoute *sensible*, sans être exclusive des autres formes d'écoute, exige tout de même qu'on « lâche » toute opération analytique, donc l'ensemble des écoutes signalées précédemment.

- Les professionnels de la musique développent des écoutes expertes très spécialisées. Ainsi les *professeurs* (d'instrument ou de voix) établissent un lien direct entre les gestes producteurs de son et les différences de qualités perçues. Ils « voient » et ressentent intérieurement sur la base de leur expérience proprioceptive, comment le son a été produit. Les *facteurs d'instrument* écoutent le son en rapport avec ce qu'ils ont éprouvé des capacités expressives des instruments issus de leur expérience de fabrication. Ils peuvent apprécier très précisément les paramètres des qualités sonores de l'attaque, de la dynamique, de la puissance rayonnée et des possibilités de variations de timbre, etc. *Les preneurs de sons* développent une écoute majoritairement centrée sur le signal et constituent leur catégorisation sonore à l'aide d'images ou de représentations mentales associées à leur pratique, ce qui leur permet d'anticiper les choix à faire au moment de régler la disposition des instruments et des microphones. Ils développent des capacités fines d'appréciation de l'espace sonore sur la base d'indices spécifiques (bruits, transitoire) afin de reproduire au mieux les effets de distance, de fusion des sons et la distribution spatiale de la scène sonore.

Nous avons peu traité de la polymodalité sensorielle et en particulier des relations étroites qui existent entre vision et audition. Il faut bien reconnaître que pour la majorité des auditeurs la vision mobilise la plus grande partie de l'attention disponible. John Blacking[54] écrit que « les musiciens savent qu'il est possible de faire passer une exécution mauvaise ou incorrecte quand on est devant un public qui regarde mais n'écoute pas », et chacun sait que fermer les yeux favorise une écoute attentive. Si un conflit d'interprétation surgit lors de la coïncidence temporelle de deux événements hétérogènes, l'un sonore et l'autre visuel, c'est ce dernier qui pèse dans la construction du sens[55]. Il est vraisemblable que la constitution en mémoire des catégories sonores inclut de fortes associations entre formes temporelles visuelles (articulations, gestes, jeu instrumental) et formes acoustiques.

Hormis les exercices d'écoute ou les tests d'acoustique, l'auditeur est généralement libre de décider de l'échelle temporelle d'écoute. Du transitoire d'attaque à la structure musicale, celle-ci varie de la milliseconde à la minute, ou plus. Les différents niveaux d'appréhension d'une musique peuvent ainsi être explorés de façon infinie, autant de fois que se renouvellent les motivations et au fur à mesure que la familiarité rend plus prévisibles les structures cachées ou complexes.

En guise de cadence finale à ces réflexions, voici, parmi bien d'autres, trois extraits d'œuvres musicales (environ 1 min), caractéristiques pour leurs « formes typiques », à écouter en regardant leurs images sonagraphiques (figures 4.45 1, 2 et 3) ou en cliquant dessus dans la version PDF de l'ouvrage, fournie sur le DVD-Rom d'accompagnement.

Son 4.40 (64'')

Tambour (improvisation au pakhavaj)

1/ Son **4.40** : une séquence musicale rythmique improvisée au *pakhavaj*, tambour indien qui permet plusieurs types de frappes reconnaissables par leur hauteur spectrale et leur durée. L'organisation temporelle s'inscrit dans le cadre d'un rythme cyclique de 12 battues. Au cours du premier cycle le musicien « dit » les sons par onomatopée. Chaque début de cycle est repéré sur le sonagramme par un curseur rouge. L'analyse est limitée à 1,5 kHz.

54. *Le sens musical*, 1973, p. 19.
55. Voir la démonstration de McGurk, séquence vidéo téléchargeable à partir de *Wikipedia*. (Entrée « Effet McGurk », vue en 2013).

2/ Son **4.41** : une séquence de musique polyphonique jouée sur un pianoforte (diapason la_3 = 430 Hz). Les blocs d'accords du début et les figures mélodiques (arpèges, notes répétées, traits rapides descendants) sont caractéristiques de l'écriture musicale pour cet instrument. Le choix de la bande passante, limitée à 4 kHz, zone qui correspond à celle des fondamentaux musicaux, est un compromis entre la lisibilité des hauteurs et le rendu du spectre. La note la plus aiguë du passage (fa_6 = 2636 Hz), repérée par un astérisque sur la figure, est le début de l'arpège descendant. La netteté et la lisibilité de l'image sont caractéristiques du son de l'instrument, mais aussi du jeu de l'interprète.

Son 4.41 (72″)

Pianoforte (extrait d'une polonaise de Chopin)

3/ Son **4.42** : une séquence de musique créée à l'ordinateur (1969). L'analyse s'étend à 12 kHz, car dans cette œuvre, la structure acoustique des sons (qui occupent tout le spectre) fait partie intégrante de la composition musicale. pour ce style d'œuvre, la représentation sonagraphique fonctionne comme une partition musicale.

En 1984 Robert Cogan a publié dans *New images of musical sound*, les premières analyses spectrographiques visualisant la structure sonore des musiques.

Son 4.42 (42″)

Synthèse à l'ordinateur (extrait de Mutations, *de J.-C. Risset)*

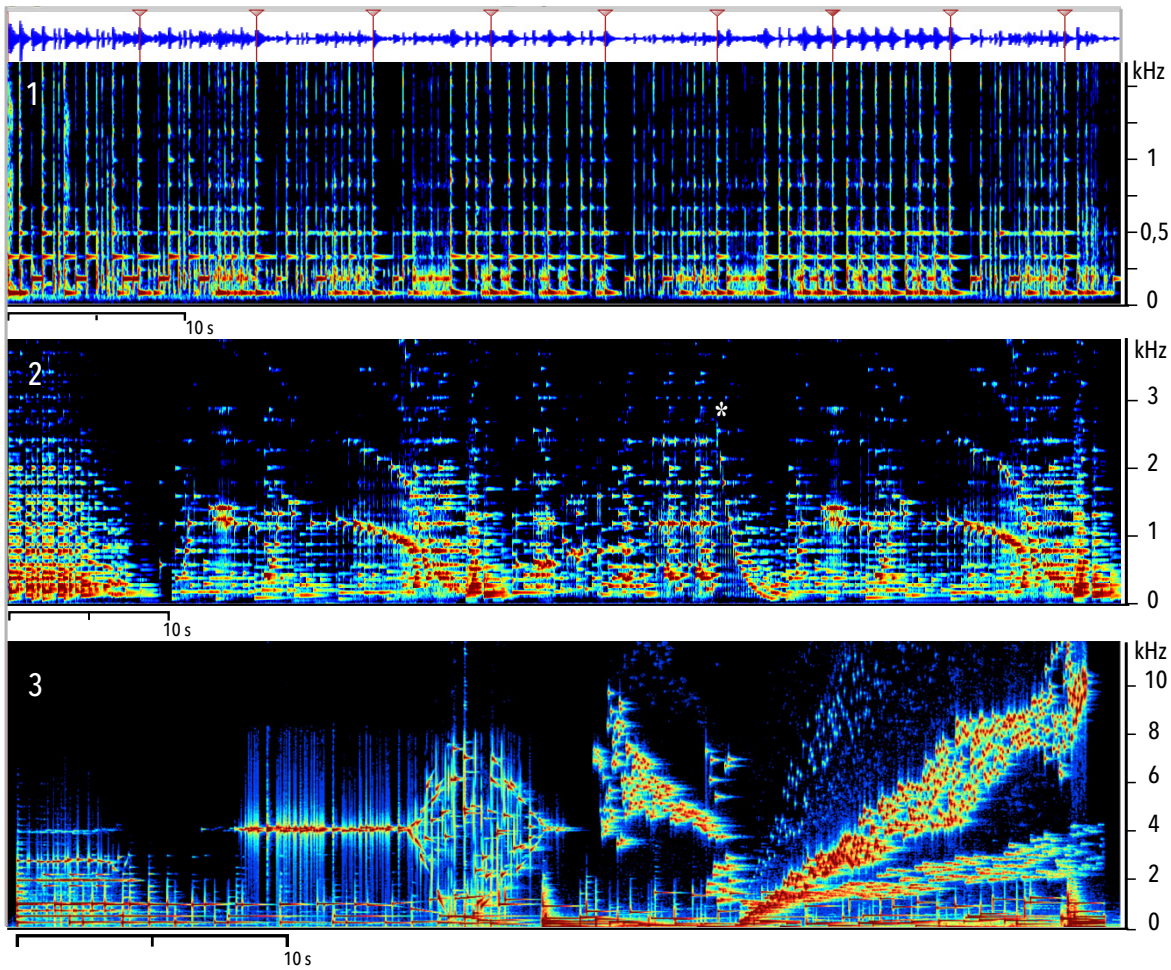


Figure 4.45 Analyses sonagraphiques de trois extraits de musiques évoquant trois modes d'écoute. **1/** Musique rythmique concentrée dans les basses fréquences, de structure essentiellement temporelle et spectrale. Analyse limitée à 1,5 kHz. **2/** Musique pour piano évoluant dans la zone des « fondamentaux musicaux ». L'écoute s'attache simultanément aux rapports de hauteurs mélodiques et aux rapports de durées. Analyse limitée à 4 kHz. **3/** Musique composée à l'ordinateur occupant un large spectre. Tous les modes d'écoute sont sollicités, dont l'écoute spectrale. Analyse de la voie droite, limitée en fréquence à 12 kHz.

5. Conclusions

L'auditeur au centre des processus de perception sonore

Nous avons proposé un parcours de découverte des différents aspects de la perception sonore en nous fondant sur des exemples extraits du monde réel. Cette approche, qui vient compléter les recherches sur la perception humaine plus généralement orientées sur la vision, s'inscrit dans le courant « écologique », lequel stipule que percevoir est l'acte d'un vivant qui donne sens au monde. L'auditeur, qui est au centre du processus, est toujours « situé » : dans un lieu, à un moment donné de son existence, plongé dans une activité et en interaction constante avec les autres membres du groupe humain auquel il appartient. Tous les sens participent d'une telle perception et en particulier les mouvements de celui qui écoute ou qui produit lui-même les sons.

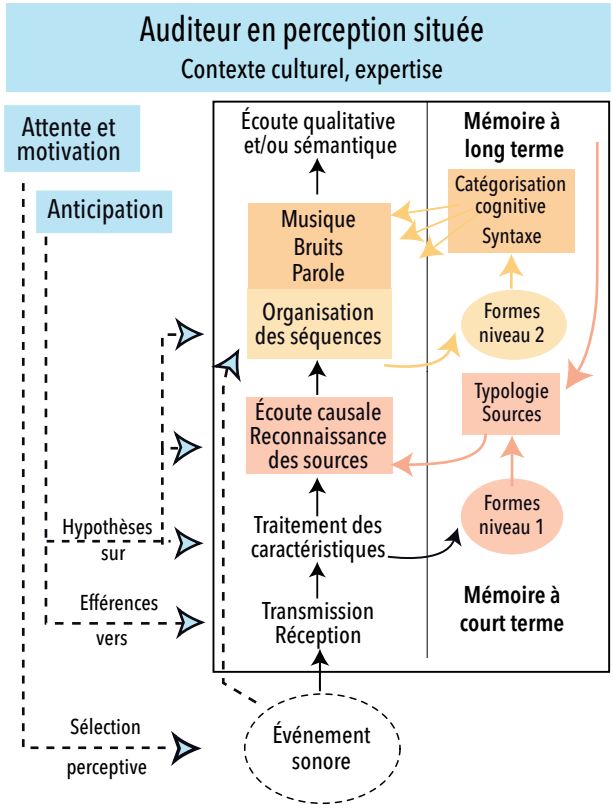


Figure 4.46 Proposition d'un schéma illustrant les principales étapes du traitement perceptif d'événements sonores. Voir aussi figure.4.3, page 147 et figure 4.18, page 165.

Telle que nous l'avons présentée, la perception suppose donc une interaction constante entre des flux descendants initiés par l'auditeur qui sélectionne ce qui est pertinent pour lui et les flux ascendants des données sensorielles qu'il traite et interprète[56]. Mû par ses motivations personnelles et se fondant sur la prévisibilité des événements sonores habituellement corrélés à une situation donnée, l'auditeur procède le plus souvent par anticipation et vérification des formes sonores déjà catégorisées en mémoire (figure 4.46). Il oscille entre deux modalités perceptives

56. Les recherches actuelles en neurosciences commencent à mettre en évidence le rôle actif de l'auditeur qui agit, *via* les circuits efférents (descendants), sur les modalités de réception et de prétraitement des données sensorielles. Voir Shamma S. et Micheyl C., 2011.

antagonistes et complémentaires : l'une, causale, qui vise l'identité des sources sonores, et l'autre, qualitative, qu'il peut mettre en œuvre à tout moment selon ses choix, et selon les modes de communication sonore dans lesquels il s'engage.

Deux stratégies d'écoute

Nous avons distingué deux niveaux de structuration de la perception sonore. Le premier, directement en prise avec les caractéristiques initiales de la stimulation acoustique fournit dans un temps très court les éléments de décision à l'écoute causale dont l'enjeu est vital. Le second, au cours duquel des opérations complexes d'analyse et d'évaluation font appel aux fonctions cognitives de haut niveau, nous permet d'appréhender l'organisation des séquences sonores de durée plus longue, telles que la parole et la musique (voir figure 4.47).

Quoi ? Où ?	Identification	Écoute causale	Typologie acoustique
Comment ?	Qualification	Écoute qualitative	Prototypes personnels

Figure 4.47 Schéma des deux types d'écoute : causale et qualitative.

Premier niveau d'écoute : les sources sonores

Le son est incontestablement un phénomène matériel dont le déroulement temporel porte la trace des événements qui en ont été la cause. Au premier niveau s'opère la reconnaissance des sources sonores sur la base des indices temporels et spectraux correspondant aux actions de mise en vibration et aux réponses acoustiques de la structure ébranlée. En nous basant sur les principales co-occurrences spectrotemporelles nous avons proposé (figure 4.43, page 188) une typologie acoustique élémentaire des sources sonores mécaniques usuelles. Celle-ci se fonde sur trois distinctions fondamentales : le type d'excitation qui conditionne le transitoire d'attaque (percussif ou non), la présence ou l'absence de fréquences discrètes immédiatement après le début de l'excitation, le degré de stabilité de ces fréquences, habituellement corrélé avec le fait que l'émetteur est vivant ou mécanique. L'élaboration individuelle d'une typologie sonore de niveau 1, renforcée par l'expérience quotidienne des sons que nous produisons par nos propres mouvements, constitue une grille d'entrée efficace pour discriminer les formes types et reconnaître rapidement la source du son. Ainsi présentée, l'écoute causale, ou écoute de vigilance, est une prise d'information immédiate sur les sources sonores s'effectuant dans une durée très brève de quelques dizaines de millisecondes. Véritable porte d'entrée de la perception sonore, elle participe de tous les modes d'écoute, parole et musique comprises. L'excellente aptitude à l'écoute causale dont font preuve les auditeurs de toutes cultures atteste de l'extraordinaire capacité des êtres vivants à repérer et à extraire les invariants et les régularités de structure à travers la variabilité des réalisations sonores réelles.

Deuxième niveau d'écoute : les séquences

La grande majorité des sons qui nous concernent quotidiennement sont des signaux de communication que nous avons appelés de second niveau, qui, contextualisés, prennent sens comme « parole », « bruit », « musique ». Il s'agit d'un changement radical de la perception sonore puisque l'information n'est plus directement portée par la source, mais par les variations spécifiques qui lui ont été appliquées : ce qui est transmis est codé. Un signal d'alarme, un ordre, l'énoncé d'un concept, l'incipit d'une chanson sont des séquences sonores symboliques dont la structure sémantique

s'organise sur une échelle temporelle supérieure à 0,5 s, pouvant aller jusqu'à 20 s, voire plus en musique. Le traitement de ces formes de second niveau nécessite donc des capacités de rétention en mémoire immédiate beaucoup plus considérables.

Les formes spectrotemporelles les plus fortes émergent d'un ensemble sonore confus, supportent les anamorphoses et se prêtent à des jeux d'ambiguïté perceptive recherchés dans les arts sonores. Chaque individu se constitue ainsi un répertoire de formes sonores dont le contenu et l'organisation évoluent au fur et à mesure de l'acquisition des connaissances et se stabilise lors des échanges verbaux avec les autres membres de la communauté.

Formes et catégories

Si nous percevons des formes, c'est par la catégorisation que nous donnons sens au monde (Dubois, 1993). La musique n'existe pas sans la matière sonore des sources instrumentales qui la produisent. Or, les notions d'acoustique musicale sont trop souvent présentées comme une correspondance évidente entre des paramètres physiques et les éléments du solfège. La notion de hauteur musicale, par exemple, n'est pas seulement liée à la saisie de la fréquence d'un son : elle doit être repérée, stabilisée indépendamment des autres variations sonores (intensité, couleur timbrale) – ce que favorise la pratique d'instruments différents – et sans cesse renouvelée en mémoire. Les concepts de la catégorisation prototypique éclairent ces processus et permettent de mieux comprendre la richesse et la diversité des écoutes musicales.

Les relations entre les notions de « formes » et de « catégories » ne sont pas aisées à saisir, car elles sont exposées dans des ouvrages de disciplines différentes. Comme nous le présentons dans la figure 4.48, les formes sont généralement attribuées aux objets du monde matériel alors que les catégories, traitées dans les ouvrages de philosophie et de cognition, ont plutôt trait au monde des idées, et plus particulièrement à leur expression dans le langage. Pour les deux notions, l'enjeu est d'extraire des caractéristiques en comparant les similitudes et les différences des formes – ou des catégories. Pourtant, la notion de forme est fondamentalement globale alors que les catégories peuvent aussi être définies par une liste de propriétés.

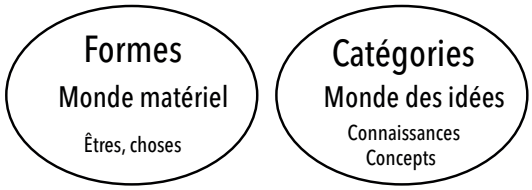


Figure 4.48 Formes ou catégories ?

Problème de la variabilité

Le problème le plus fascinant de la perception humaine et, plus généralement, de tous les êtres les vivants, est celui du traitement de la variabilité intrinsèque au monde naturel. Excepté les signaux d'alarme stéréotypés, les séquences de parole, de musique et les sons de l'environnement supportent d'incessantes transformations que la Gestalttheorie nous a permis d'appréhender.

La perception causale des sources et la perception sémantique de la parole ne réussissent que lorsque nous sommes parvenus à ignorer les variations contingentes pour appréhender efficacement les structures spectrotemporelles qui font sens.

À l'opposé, la perception qualitative de la musique, celle de la voix qui parle ou celle du bruit de la pluie, sont des modalités d'écoute de type analytique et discriminant, par lesquelles nous prenons justement plaisir à saisir les plus infimes variations des sons, ce qui les rend intéressants à écouter, et comme l'écrit si joliment Mersenne, « nous les fait admirer et souhaiter ».

> *Choisissez tel son que vous voudrez, et l'oyez continuellement, il vous endormira, ou vous fera mal à la teste. Le son d'une fluste posée sur un sommier, estant continu, est merveillement importun, et déplaisant ; et celui d'un luth le serait encore d'avantage, s'il se pouvait continuer aussi lontemps que celuy de la fluste.* C'est donc la varieté qui rend le son agreable *et s'il n'est varié, il merite plustost d'estre appellé bruit que son harmonique ; et parce que la varieté ne se peut faire qu'avec le temps, tous les sons ont besoin de temps pour faire quelque varieté, et quelque impression dans l'esprit qui nous les fait admirer et souhaitter.*
>
> Mersenne, 1636, *Livre premier des instruments, Proposition : « À savoir quel est le plus agréable son de tous les Instruments de Musique »*

De ces deux stratégies d'écoute, totalement antagonistes, la première est indispensable à la survie. La seconde se déploie sous des formes les plus diverses et ne cesse de se transformer au long de la vie : chez les passionnés du son elle est, selon l'expression de Blacking, la source inépuisable de l'« écoute créatrice ».

> *On néglige trop souvent l'importance de l'écoute créatrice dans les discussions sur l'aptitude musicale, alors qu'elle est aussi fondamentale pour la musique qu'elle l'est pour le langage.*
>
> Blacking, J., 1973, p. 18.

6. Documentation sonore

1/ *Auditory Scene Analysis* – Disque de 43 démonstrations réalisées par Albert Bregman et Pierre Ahad (1995).

Voir la page personnelle d'A. Bregman où l'on peut entendre en ligne ces démonstrations, et lire les commentaires de l'auteur. Les exemples sonores au format MP3 peuvent être téléchargés à partir du site, ainsi que le livret d'explications. Le CD au format 16 bits PCM peut-être commandé au MIT Press :

<http://webpages.mcgill.ca/staff/Group2/abregm1/web/index.htm>
(visitée en août 2014)

2/ *Musical illusions and paradoxes* – Disque de 23 exemples réalisés par Diana Deutsch.

Voir la page personnelle de Diana Deutsch qui présente une documentation (écrite et sonore) très fournie :

<http://deutsch.ucsd.edu/psychology/pages.php?i=101>.

Voir le site : <http://philomel.com/musical_illusions> (visité en août 2014), où l'on peut entendre quelques exemples sonores.

7. Les sons du chapitre 4

7.1. Expériences d'écoute

Son 4.1 – La consigne d'écoute est : *décrivez les qualités sonores de l'instrument*. Il s'agit de prendre conscience des difficultés que l'on éprouve à percevoir les qualités d'un son qu'on ne peut pas rattacher à une catégorie connue. [M. C.]

Son 4.2 – Combien de temps vous a-t-il fallu pour reconnaître l'instrument ? Début de la phrase (7"), puis la phrase entière. A. Lagoya, 1965. [Archives LAM, M. C.]

Son 4.3 – Écoute d'une scène sonore ordinaire : faites la liste de tous les sons que vous entendez. [M. C.]

7.2. Formes sonores (sources)

Son 4.4 – Collection A de 29 sons brefs issus de sources acoustiques connues. Vous pouvez aisément nommer chacun d'eux. En voici la liste. 1) démarrage d'aspirateur ; 2) débouchage (bouteille de champagne) ; 3) bris de verre ; 4) aboiement ; 5) cloche ; 6) corbeau ; 7) cri de bébé ; 8) cymbale ; 9) débouchage (bouteille de vin) ; 10) démarreur de voiture ; 11) écoulement d'eau ; 12) tuyau d'orgue ; 13) crissement de pneus ; 14) cri de goéland ; 15) goutte d'eau ; 16) couinement de porte ; 17) klaxon ; 18) machine à écrire ; 19) marteau piqueur ; 20) bêlement de mouton ; 21) rire ; 22) sifflet à roulette ; 23) sifflet de train ; 24) sonnette de porte ; 25) sonnerie de téléphone ; 26) toux ; 27) violon ; 28) chant : « o » ; 29) insecte volant.

Son 4.5 – Collection B de 29 sons brefs issus de l'activité quotidienne. Ces sons n'ont pas de nom spécifique, mais vous pouvez probablement décrire leur mode de production. En voici la liste. 1) déchirement (papier) ; 2) cliquetis (cutter) ; 3) frottement (chiffon sur mur) ; 4) grattement (ongles sur le tissu du matelas) ; 5) froissement (papier) ; 6) frottement de deux tissus ; 7) grattement (couteau) sur un tamis métallique ; 8) déchirement (tissu de toile) ; 10) maniement de papier journal ; 11) raclement avec cuillère de bois ; 12) froissement de papier cellophane ; 13) souffle sur fleurs séchées ; 14) vaporisateur à eau ; 15) grattage (radis) ; 16) souffle dans tuyau de flûte ; 17) écrasement d'un morceau de pain ; 18) maniement de papier aluminium ; 19) pain que l'on coupe (couteau à scie) ; 20) eau qui bout ; 21) arrachement d'une feuille de papier (cahier à spirale) ; 22) frottement de lime (ongles) ; 23) friture (poêle) ; 24) grincement (chaise) ; 25) grincement (porte de placard) ; 26) grincement (table) ; 27) épluchage (pomme) ; 28) grincement (chaise) ; 29) pschit d'une bombe aérosol. [M. C.]

Son 4.6 – Son « impossible ». Un son de cymbale percutée dont le défilement temporel est inversé. [M. C.]

Son 4.7 – Sons ambigus. La structure temporelle de ces sons, entre percussion et entretien, désoriente l'écoute des auditeurs qui ne connaissent pas l'instrument. Beaucoup suggèrent une production synthétique. [P. Gaillard]

Son 4.8 – Son déstructuré temporellement. La fréquence fondamentale d'un son de violon a été supprimée par filtrage au début du son puis elle apparaît vers 1,2 s. Elle est alors perçue comme un son autonome, étranger au son de violon dont elle fait pourtant intrinsèquement partie. [M. C.]

Son 4.9 – Le son de violon original (la_3) de l'expérience précédente. Son joué sur la corde à vide. [M. C.]

Son 4.10 – Trois signaux d'avertissement : un sifflet simple, un sifflet à roulette et une sonnette de vélo. Corsin Vogel, 1999, thèse. [Archives LAM]

7.3. Formes sonores : séquences

Son 4.11 – Trois séquences de signaux de priorité automobile : pompiers, police, ambulance. [*Ibid.*]

Son 4.12 – Le mot « aujourd'hui » dit avec deux voix différentes : voix ordinaire puis voix chuchotée. [M. C.]

Son 4.13 – Le mot « aujourd'hui » articulé avec le son d'une guimbarde puis synthétisé à l'icophone. [M. C.]

Son 4.14 – Mélodie jouée au piano. Pierre Schaeffer, *Solfège de l'objet sonore*, 1967-1998, CD I, n° 9b. [Ina-GRM[57]]

Son 4.15 – Même mélodie chantée, *ibid.* n° 9a. [*Ibid.*]

Son 4.16 – Même mélodie jouée au marimba, *ibid.* n° 8. [*Ibid.*]

Son 4.17 – Trois séquences sonores familières. 1) Bruit domestique : ouverture d'un placard (fermeture aimantée) dont la porte grince, saisie d'une bouteille puis fermeture du placard. 2) Phrase dite par une voix féminine « j'ai mangé tous les champignons ». 3) Mélodie jouée à la flûte traversière. [M. C.]

Son 4.18 – Son 4.18a : séquence de six gouttes d'eau. Son 4.18b : séquence de six aboiements de chiens. [M. C.]

Son 4.19 – Anamorphose fréquentielle de voix parlée. Voix chuchotée synthétisée à l'icophone. De la première à la quatrième présentation, le dessin, support de la forme acoustique de la phrase synthétisée, a été anamorphosé verticalement par rapport à l'original, dans les proportions suivantes : 1,5 ; 1,25 ; 1 (original) ; 0,75. Entre le premier et le dernier exemple, toutes les fréquences sont divisées par deux. Son 4.19a : la phrase entière « le petit chat fait sa toilette » avec les quatre anamorphoses. Son 4.19b : le mot « chat » extrait des quatre phrases. Son 4.19c : le mot « toilette » extrait des quatre phrases. Icophone 02 du LAM, 1968. [M. C.]

Son 4.20 – La petite scène sonore déjà présentée dans le Son 4.3. [M. C.]

Son 4.21 – Musique polyrythmique, trois séquences. Son 4.21a : extrait du début de la pièce (9'') correspondant à l'analyse sonagraphique de la figure 4.26. Son 4.21b : montage de transformations sonores avec Audiosculpt ; 1) très court extrait ; 2) le même sans le tambour en sablier ; 3) le tambour en sablier extrait et entendu isolément ; 4) l'extrait complet du début. Son 4.21c : la pièce dans sa totalité. Musique de cérémonie enregistrée au Sud-Viêtnam, 1963, Trân Van Khe. [Archives LAM]

Son 4.22 – Court extrait de *Quatuor II* de Betsy Jolas. Disque 33 t. La voix de son maître, 1964. [B. Jolas[58]]

7.4. Formes et flux sonores : groupements et fissions

Son 4.23 – Structuration spontanée d'une séquence de sons. Un son synthétique de hauteur et d'intensité constantes est répété régulièrement avec d'infimes variations spectrales, totalement aléatoires. Pourtant, la plupart des auditeurs « perçoivent » spontanément un rythme qui n'est pas présent dans le son. *Gluing tones,* édité par J. Sundberg, 1992, CD, n° 1. [G. Bennett]

Son 4.24 – Incidence du tempo et de l'intervalle mélodique entre deux flux. Une séquence de sons A, B, A est présentée à des tempos différents. Son 4.24a :

57. Reproduit avec l'aimable autorisation de Ina-GRM.
58. Voir http://www.betsyjolas.com/niv_2.php3?ch=6&nav=0. Ouvrir Discographie/Écouter. L'extrait se situe 1' 48'' après le début (consulté en 2015).

lorsque l'intervalle en fréquence A-B est grand (1400-500 Hz, soit une octave et une quinte diminuée) et que la cadence est lente, on entend généralement un mouvement mélodique A, B, A. Si la cadence s'accélère, deux flux se séparent : AAA et B intermittent. Son 4.24b : le phénomène ne se produit pas pour un intervalle A-B faible (1400-1320 Hz, soit un demi-ton). *Demonstrations of Auditory Scene Analysis*: The Perceptual Organisation of Sound. CD, n° 3, M.I.T. Press,1995. [A. Bregman et P. Ahad]

Son 4.25 – Flux mélodiques et rythmes variables. Deux mélodies sont en présence : la première (flux 1) est constituée par un son pur dont la fréquence descend graduellement de 2000 à 500 Hz puis remonte à 2000 Hz ; la seconde (flux 2) est constituée d'un son pur de 1000 Hz. La perception change en fonction de l'écart entre les sons des deux mélodies, en particulier au moment du croisement. La durée de chaque son isolé est de 60 ms. Entre deux sons successifs du flux 1, l'intervalle temporel Δt = 200 ms ; entre deux sons successifs du flux 2, Δt = 400 ms ; le flux 2 (fixe) démarre 100 ms après le flux 1. Van Noorden L. P., 1975, thèse. [Archives LAM]

7.5. Ségrégations (spectrale, spatiale)

Son 4.26 – Mélodie « arlequin » réalisée par montage de sons issus d'instruments différents. [M. C.]

Son 4.27 – Dans cet extrait de l'*Offrande musicale* de J.-S. Bach orchestrée par Anton Webern, l'auditeur est-il plus enclin à suivre la continuité instrumentale ou la continuité mélodique ? Anton Webern, *Œuvres complètes*, Sony, CD 1, n° 1. [SM3K 45845)]

Son 4.28 – Fission timbrale. Une petite séquence rapide de notes disjointes sans rythme est présentée de trois façons différentes. Son 4.28a : mélodie synthétisée avec un seul timbre instrumental. Son 4.28b : mélodie synthétisée avec une alternance de deux timbres, ce qui induit une fission mélodique. Son 4.28c : autre exemple plus contrasté. [D. Wessel]

Son 4.29 – Deux mélodies entrelacées jouées au clavecin sur un seul clavier, jeu legato. Les deux parties supérieures de la Variation Goldberg n° 3 de J.-S. Bach, mesure 8. [M. C.]

Son 4.30 – (à écouter en stéréo ou au casque). Même exemple musical que précédemment, joué sur deux claviers séparés. [M. C.]

Son 4.31 – (à écouter sur enceintes). Expérience portant sur le début du finale de la *Symphonie n° 6* de Tchaikovsky. Son 4.31a : la partie mélodique des premiers violons. Son 4.31b : la partie mélodique des deuxièmes violons. Son 4.31c : les deux parties en écoute dichotique : une partie sur chaque canal (séparables au casque). Son 4.31d : extrait orchestral du début du finale : University of Chicago Symphony Orchestra, 2010. [Sons a-b-c, source inconnue ; Son d : *archive.org.details/uso20100130*]

Son 4.32 – (à écouter sur enceintes). Ségrégation spatiale de deux mélodies entrelacées : même timbre, même tessiture, homorythmie alternée. Dans l'exemple 4.32a, les deux parties, mélangées dans les canaux gauche et droite, sont reproduites en monophonie. On entend successivement la mélodie 1 seule, puis avec la mélodie 2 ; ensuite la mélodie 2 seule ; et de nouveau les deux mélodies ensemble. Dans l'exemple 4.32b, les deux parties, toujours mélangées, sont progressivement séparées et distribuées chacune sur un canal différent. Le suivi individuel de chaque partie s'éclaircit et devient possible sans effort. Musique intitulée *Ssematimba ne Kikwabanga* enregistrée au xylophone *amadinda*

(Ouganda) et numérisée par U. Wegner[59]. *Demonstrations of Auditory Scene Analysis*, CD, n° 41, 1995. [A. Bregman[60] & P. Ahad]

Son 4.33 – Transcription sonore de formes visuelles : synthèse à l'icophone, 1969. Le Son 4.33a correspond au dessin : « une montagne avec de la pluie ». Il s'agit ensuite d'attribuer les sons 4.33b, 4.33c, 4.33d, 4.33e à chacun des quatre dessins de la marge. Les réponses sont : son b (2), son c (4), son d (3), son e (1). Remarquez combien la reconnaissance d'un mot parlé (aujourd'hui, Son 4.33d) est rapide et s'impose de façon globale. [M. C.]

Son 4.34 – Forme vocale ambiguë. Vibrato ou trille ? [M. C.]

Son 4.35 – L'exemple précédent intégré dans deux contextes mélodiques différents. Son 4.35a : l'extrait s'entend comme un vibrato. Son 4.35b : l'extrait s'entend comme un trille. Pour plus de détails, se reporter au § 4.4 du chapitre. [M. C.]

7.6. Anamorphoses sonores

Son 4.36 – Cinq étapes de la transformation d'une forme sonore en une autre (*morphing*) : du piano à la contrebasse. *Formalization and Epistemology*, 1993, thèse. [I. Godøy]

Son 4.37 – Exemple d'une anamorphose temporelle (ralentissement de la succession des impulsions vocales) produisant un saut catégoriel surprenant : de la voix humaine à la clochette. 1980. [X. Rodet]

7.7. Catégorisation (voyelle, timbre)

Son 4.38 – Deux sons brefs chantés successivement. Vous entendez probablement les voyelles « a » puis « o ». [M. C.]

Son 4.39 – La deuxième voyelle « o » suivie de la phrase musicale complète *Ave Maria* ; puis de même, la première voyelle « a » suivie de la phrase complète *Ave Maria*. Vous pouvez maintenant écouter ces deux phrases chantées en portant votre attention sur les différences de qualité vocale (son couvert pour le premier et son ouvert pour le deuxième). Chant : B. Chuberre ; thèse N. Henrich, 2003. [Archives LAM]

7.8. Formes et musiques

Son 4.40 – Musique de l'Inde jouée sur un *pakhavaj* (tambour à deux peaux), extrait (60'') du début d'une improvisation par Raja Chatrapati Singh[61]. [Archives LAM]

Son 4.41 – Musique pour pianoforte. Extrait (0'33 à 1'45) de la *Polonaise en sol mineur* kk IVa/3 ; CD *Frédéric Chopin*, n°1, Pierre Goy. [Cantando 2517]

Son 4.42 – Musique composée à l'ordinateur. *Mutation,* extrait (4'38 à 5'19) ; CD C 1003, n° 4. Jean-Claude Risset, 1969. [Ina-GRM]

59. Wegner, U., 1990, *Xylophonmusik aus Buganda (Ostafrika). Musikbogen 1. Wege zum Verständnis fremder Musikkulturen.* Wilhelmshaven, Florian Noetzel.
60. Reproduit avec l'aimable autorisation de l'auteur.
61. Voir le site http://en.wikipedia.org/wiki/Pakhavaj.

CHAPITRE 5

PERCEPTION DES QUALITÉS SONORES : L'INTENSITÉ

1. Introduction

Percevoir les variations d'intensité sonore et les apprécier ne nécessite pas d'apprentissage particulier. C'est une donnée immédiate de la sensation auditive, une compétence vitale puisque l'intensité est liée à l'énergie mise en jeu à la production du son : un son très fort nous met en alerte avant même que nous prenions soin de l'écouter, car il peut être signe de danger.

Dès qu'il devient nécessaire de préciser la notion d'intensité, et en particulier de la mesurer, plusieurs problèmes se posent. Même si l'on connaît l'énergie mise en jeu à la production d'un son, les modalités de dissipation dans la structure vibrante et au cours de la propagation sont si complexes que seuls des acousticiens spécialistes de l'intensité sonore peuvent maîtriser les conditions expérimentales pour effectuer des mesures. Contrairement à la fréquence vibratoire, qui peut franchir de grandes distances sans être altérée, l'amplitude vibratoire est affectée par un grand nombre de paramètres : elle décroît au cours de la propagation et varie d'un point à l'autre de l'espace en raison des obstacles et des multiples réflexions qui se trouvent sur le parcours des ondes.

Il n'est donc pas possible, sur la base des données perçues localement, de remonter à l'intensité d'une source émettrice, d'autant que nous adaptons continuellement la chaîne ossiculaire pour capter les vibrations dans les meilleures conditions de confort. De plus, l'oreille humaine est inégalement sensible aux différentes fréquences vibratoires (voir chapitre 3, § 2.1).

La mémorisation quantitative de l'intensité sonore ne peut donc être qu'approximative et, d'ailleurs, il n'existe pas d'oreille absolue pour l'intensité sonore ! En musique, les catégories d'intensité sonore ne sont pas très précises : très fort, fort, moyen, faible, très faible. À l'opposé, les petites variations temporelles d'intensité sont perçues avec une grande finesse. Déjà performante en laboratoire avec des sons purs (chapitre 3, § 2.5), l'oreille se révèle d'une extrême sensibilité pour la discrimination d'intensité des sons complexes, lors du décryptage des sons de l'environnement sonore, pour situer une source dans l'espace ou encore pour apprécier l'expressivité musicale. Pourtant, si l'amplitude sonore est bien le paramètre physique d'entrée, l'interprétation que nous

en donnons en termes d'intensité perçue se fonde principalement sur l'interprétation d'un ensemble de transformations spectrotemporelles dues à ces variations d'intensité, et en particulier celles que nous associons à chaque catégorie de sons : une voix, un instrument de musique particulier ou tout autre phénomène sonore identifié. En jouant sur les paramètres spectraux et temporels des sons complexes pour développer une palette de sensations d'intensité d'une infinie diversité, l'expérience empirique des humains, et celle plus experte des musiciens et des facteurs d'instruments, tire admirablement parti des contraintes physiologiques que l'étude des sons purs a révélées.

2. L'intensité perçue et la zone de sensibilité de l'oreille humaine

Rappelons tout d'abord que la sensibilité de l'oreille varie avec la fréquence (chapitre 3 § 2.6). Les signaux d'alerte traditionnels exploitent à merveille cet état de fait en jouant soit sur la tessiture, soit sur la répartition de l'énergie dans le spectre.

2.1. Le rôle de la tessiture

Pour produire le maximum d'efficacité avec le minimum d'effort (c'est-à-dire chatouiller l'oreille de l'auditeur sans trop se fatiguer), le plus simple est de focaliser l'énergie dans la zone de sensibilité maximale de l'oreille humaine, soit 3 000 Hz. Il s'agit d'une zone de fréquence très aiguë, bien au-dessus des sons que la voix humaine peut atteindre. Seules quelques petites flûtes, dont le piccolo d'orchestre, jouent dans la tessiture do_6-sol_6, le plus souvent pour produire des effets « *violents et déchirants, dans un orage, par exemple, ou dans une scène d'un caractère féroce, infernal* »[1].

L'efficacité et le caractère agressif des sons de cette tessiture conviennent parfaitement au sifflet d'agent de police qui doit émerger sur le bruit de fond intense de la circulation automobile (voir figure 5.1 et Leipp, 1977, chapitre XI).

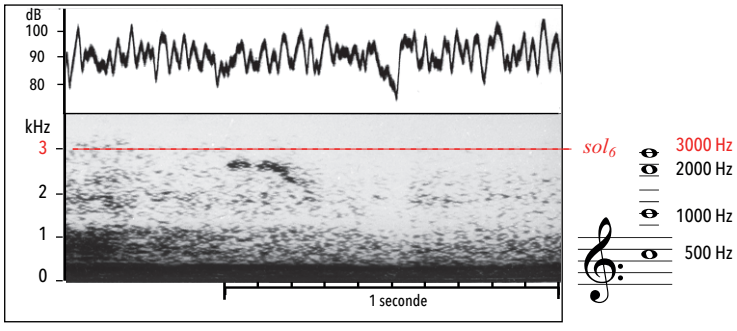


Figure 5.1 Sifflet d'agent de la circulation émergeant sur un bruit de fond intense et correspondance note-fréquence pour 3 000 Hz.
Document sonagraphique : E. Leipp (archives LAM).

Citons aussi certains instruments de musique comme le *Song Lang*[2] de la musique vietnamienne, qui joue un rôle important dans l'organisation rythmique d'un ensemble instrumental fourni, ou le piccolo de l'orchestre, seul instrument mélodique qui peut « passer au-dessus » d'une formation de plus de cent musiciens.

1. Berlioz, 1855, *Traité d'instrumentation et d'orchestration*, p. 158.
2. Voir *Bulletin du GAM* n° 55, page 7.

2.2. Les formants présents dans la zone 3 000 Hz

Pour les sons dont la tessiture est plus grave, une autre façon d'agir sur l'intensité perçue consiste à renforcer les composantes spectrales voisines de 3 000 Hz. L'exemple le plus remarquable est celui de la voix humaine. Un orateur professionnel qui parle en plein air ou qui fait face à une assemblée bruyante évite de fatiguer sa voix en criant. Il élève un peu la hauteur fondamentale et « timbre » le son pour l'enrichir en harmoniques aigus. Certains chanteurs, en particulier les chanteurs d'opéra, ont même la faculté de renforcer à volonté et de façon élective les harmoniques situés dans la zone 2 500-3 500 Hz, quelle que soit la fréquence fondamentale de la voix, en développant un formant spécifique qu'Helmholtz[3] avait déjà remarqué.

Ce formant appelé formant du chanteur (*singer formant* en anglais), présent quelles que soient les voyelles, permet à la voix d'émerger par rapport à l'orchestre et lui confère une qualité de « brillance » très appréciée. La figure 5.2 montre l'analyse sonagraphique d'une voix de baryton dans laquelle le formant, qui s'étend de part et d'autre de 3 kHz est particulièrement bien marqué (encadrement).

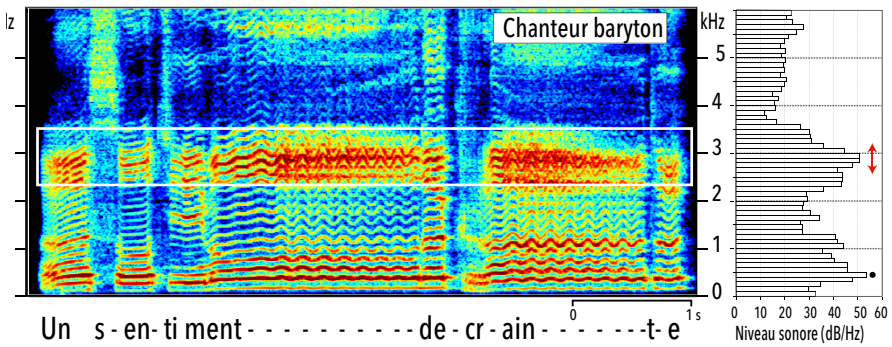


Figure 5.2 Sonagramme de la voix d'un baryton avec le formant du chanteur encadré en blanc. À droite : spectre moyenné (LTAS) montrant l'accroissement d'énergie de part et d'autre de 2 900 Hz. Cette voix « passe » au dessus de l'orchestre car le formant correspond à la zone de plus grande sensibilité de l'oreille.

Son 5.1a (6'')

Voix chantée normale

Son 5.1b (6'')

Filtrage du formant dit du chanteur

Pour comparer l'incidence du formant sur l'intensité et sur la qualité sonore, écoutez successivement le Son **5.1a** et le Son **5.1b**. Le filtrage total des harmoniques situés dans la zone du formant 2 200-3 000 Hz adoucit la voix et change la qualité timbrale de façon très importante.

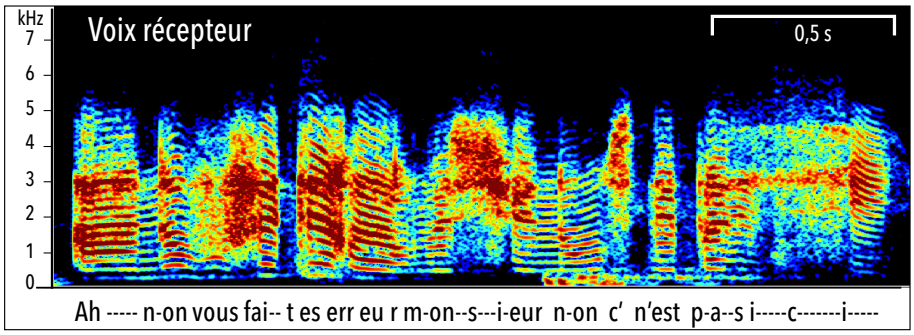
Les premiers appareils mécaniques de reproduction du son (phonographes), et encore aujourd'hui certains téléphones, présentent une particularité spectrale similaire, c'est-à-dire un renforcement des fréquences de la zone sensible de l'oreille. Parfois, la qualité sonore des signaux transmis est sérieusement distordue, et il arrive que l'on ne reconnaisse pas la voix d'une personne familière que l'on entend pour la première fois au téléphone.

3. Helmholtz, 1874, *Théorie physiologique de la musique*, p. 146.

Les deux exemples suivants, Son **5.2** et Son **5.3**, permettent de comparer la voix normale d'une locutrice et celle que reçoit l'auditeur à l'écoute de son récepteur téléphonique. Les distorsions entendues correspondent à celles qu'on voit sur la figure 5.3 : filtrage des sons graves, en particulier de la fréquence fondamentale ; coupure des fréquences aiguës supérieures à 4 500 Hz (comparez le spectre des consonnes « s, ch, z ») ; résonance marquée autour de 3 000 Hz. Le système de transmission est efficace mais peu fidèle.

Son 5.2 (4'')

Voix captée au récepteur



Son 5.3 (4'')

Voix naturelle

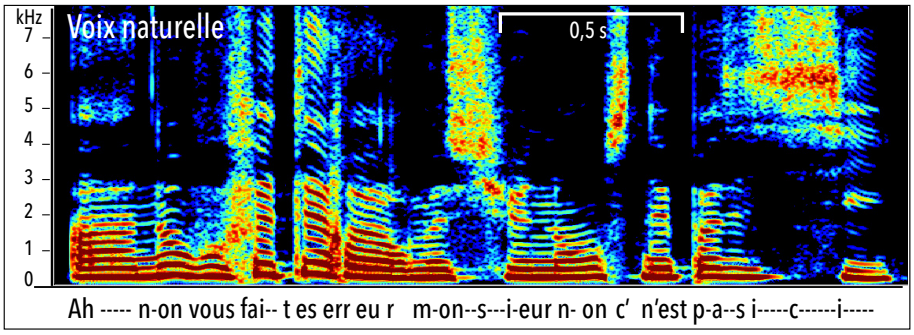


Figure 5.3 La voix au téléphone (LAM, 1986). L'enregistrement a été fait simultanément près du locuteur (en bas) et à la sortie du récepteur téléphonique (en haut). La comparaison des deux analyses fait apparaître les « distorsions » de ce téléphone ancien, le filtrage des composantes inférieures à 0,5 kHz et de celles qui dépassent 4,5 kHz. Notez l'importante accentuation des fréquences voisines de 3 kHz pour gagner en efficacité.

3. La perception de l'intensité et l'enrichissement spectral

3.1. Le crescendo musical

Les variations d'intensité sonore sont l'une des données importantes de l'expressivité musicale. Pour jouer plus fort, il faut fournir plus d'énergie : l'amplitude sonore croît. Tout naturellement les premiers instruments de synthèse ont été dotés d'une pédale de volume agissant sur le gain de sortie de l'amplificateur. Cette action, qui « grossit » le son de façon uniforme, était peu efficace et les musiciens exagéraient souvent les variations de volume pour ressentir des variations d'intensité. Dans le même temps, les analyses acoustiques d'instruments réels (Fletcher, 1934 ; Leipp, 1964) et les premières synthèses (Risset, 1969) ont mis en évidence le fait que le crescendo des instruments traditionnels impliquait non seulement une augmentation de l'amplitude globale, mais aussi d'importantes transformations spectrales.

L'exemple sonore suivant permet de prendre la mesure de cet effet. Soient deux sons ayant la même fréquence fondamentale de 500 Hz (do_4), l'un pur (Son **5.4a**) et l'autre composé de cinq harmoniques (Son **5.4b**). Tous deux ont exactement le même niveau sonore global en dB, or le son harmonique paraît nettement plus fort à l'écoute que le son pur. Toute l'énergie du premier son est concentrée à 500 Hz. Celle du second, répartie de 500 à 2 500 Hz, affecte plusieurs bandes critiques différentes (voir chapitre 3, page 98) et, de ce fait, est plus efficace à l'oreille.

Son 5.4a (2'')

Son pur de 500 Hz

Son 5.4b (2'')

Son harmonique

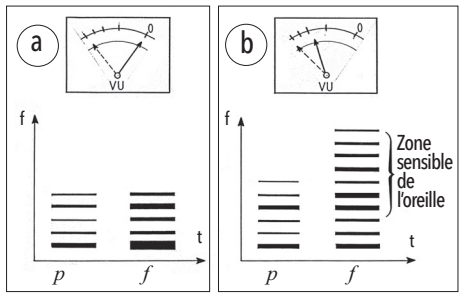


Figure 5.4 Changement d'intensité et variation spectrale. a) son synthétique joué *p* puis *f* par variation du niveau de sortie : toutes les composantes sont amplifiées dans les mêmes proportions. b) crescendo d'un instrument de musique traditionnel : enrichissement du spectre, plus particulièrement dans la zone sensible de l'oreille.

Castellengo, M., *Le livre des techniques du son*, Paris : Dunod, figure 2.5.

Lorsqu'un musicien accroît le niveau d'excitation en agissant – selon les instruments – sur la frappe, sur le débit d'air ou sur la vitesse d'archet, le son devient plus intense, mais, surtout, le spectre s'enrichit en harmoniques de rang élevé. Cet enrichissement, dû aux non-linéarités du système mécanique[4], permet de produire d'importantes variations de la dynamique perçue pour de petites variations de l'excitation. Comparez, sur la figure 5.4, le son joué *p* puis *f* avec un instrument de synthèse ordinaire (a) et avec un instrument de musique mécanique (b). C'est la synthèse par modulation de fréquence[5] qui a permis d'introduire des variations de spectre associées aux variations d'intensité. Ce paramètre (*velocity),* que le musicien peut contrôler en cours de jeu, varie avec la vitesse d'enfoncement de la touche sur les instruments à clavier.

Voici l'analyse de quelques exemples de sons musicaux réels avec des variations d'intensité. La figure 5.5 montre l'analyse spectrale d'un crescendo joué à la trompette sur une note répétée (Son **5.5**). Au fur et à mesure que l'intensité croît, le spectre s'enrichit et la qualité sonore perçue gagne en brillance (l'analyse est limitée à 10 kHz, mais les harmoniques montent au-delà).

Son 5.5 (7'')

Crescendo de trompette

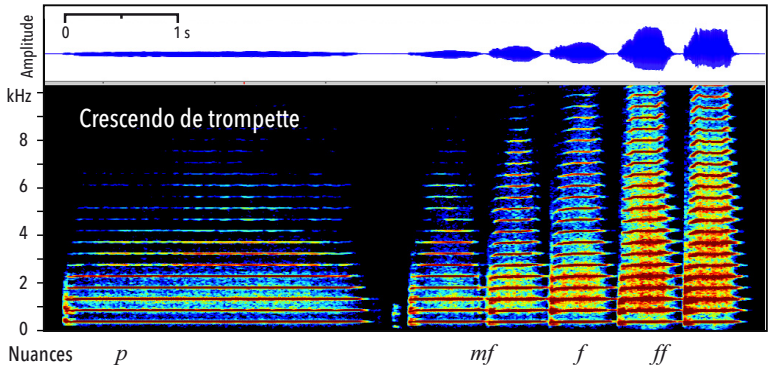


Figure 5.5 Analyse d'une note (sib_3) jouée à la trompette avec la nuance *p*, puis répétée avec une intensité croissante. Notez que l'intensité du premier harmonique (le fondamental) est peu affectée par le crescendo.

4. Si l'instrument était linéaire, l'augmentation d'amplitude de l'excitation produirait un accroissement de l'amplitude globale du spectre, sans enrichissement en nombre d'harmoniques (voir Glossaire).
5. Voir les travaux de John Chowning (1973) et leur application au synthétiseur DX7 (Yamaha).

Son 5.6 (21'')

Saxophone, son filé

L'exemple suivant (Son **5.6**), joué au saxophone alto, est celui d'un « son filé » (sib_2) qui dure 25 s. Comme pour la trompette, l'analyse de la figure 5.6 révèle l'extraordinaire complexité sonore du phénomène de crescendo : au moment du fortissimo, on compte plus de soixante harmoniques. On notera aussi le renforcement des harmoniques situés vers 3 kHz (marqué par le noircissement de l'image dans cette zone), qui se produit au moment de l'intensité maximale.

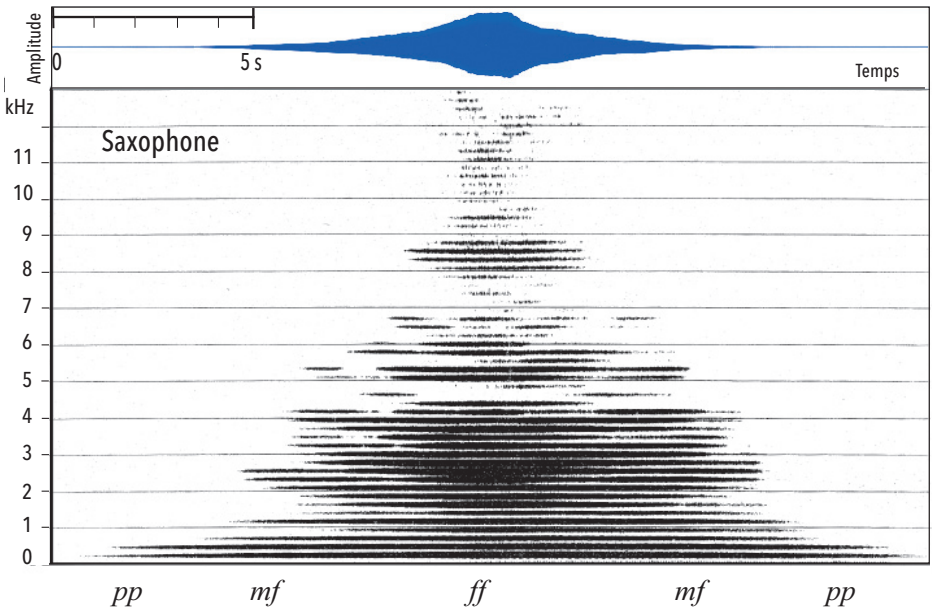


Figure 5.6 Analyse d'un « son filé » joué au saxophone alto (sib_2). En haut : courbe enveloppe du signal. En bas : sonagramme. Du *pp* au *ff*, le spectre s'enrichit de façon considérable avec une accentuation marquée de la zone des 3 kHz.

Nous avons mesuré, avec le logiciel Praat, l'écart en dB entre les sons *pp* du début et de la fin et le son *ff* (figure 5.7) : cet écart est de l'ordre de 40 dB. La beauté de la figure et son élégante symétrie témoignent de la maîtrise de l'instrumentiste.

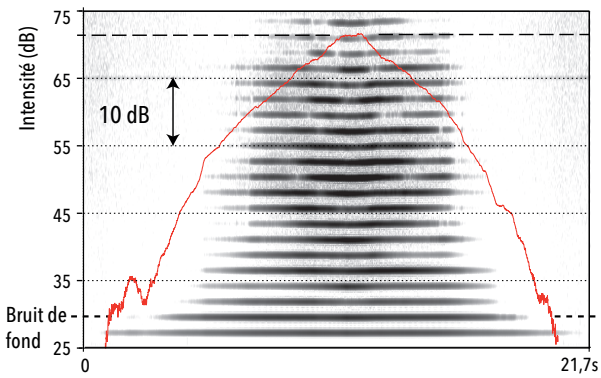


Figure 5.7 Mesure de la variation d'intensité pendant le crescendo du son de saxophone. L'écart total entre le bruit de fond et le *ff* est d'environ 40 dB.

3.2. Les sourdines, l'effet d'éloignement et d'écho

L'intensité d'une source diminue avec la distance, mais, s'il s'agit d'un son réel, un autre indice nous renseigne sur l'éloignement : l'atténuation des composantes aiguës du spectre, qui sont plus rapidement absorbées que les graves au cours de la propagation. Cet effet est exploité de diverses façons dans la facture instrumentale. Les sourdines de violon et de violoncelle, placées sur le chevalet, en modifient la masse et l'inertie, afin de décaler l'efficacité de la transmission vers les plus basses fréquences. On trouve, à l'orgue, plusieurs dispositifs pour modifier l'intensité du rayonnement sonore, en particulier dans les claviers d'écho et du récit. Les tuyaux du clavier d'écho sont situés dans le soubassement du grand buffet, de sorte que le son parvient atténué et assourdi. Ceux du clavier de récit de l'orgue romantique sont enfermés dans une boîte *expressive* comportant des volets dont l'organiste peut commander l'ouverture depuis la console pour effectuer des nuances globales. D'une façon générale, tout dispositif ayant pour effet de filtrer les hautes fréquences diminue la sensation d'intensité et suggère une impression d'éloignement.

3.3. Le problème des sons graves

La production des sons de basse fréquence pose des problèmes particuliers. D'une part, il faut des objets de grandes dimensions, souvent lourds à manier, d'autre part, l'oreille voit sa sensibilité diminuer rapidement dès que la fréquence descend au-dessous de 150 Hz, ce qui oblige à fournir de plus en plus d'énergie au fur et à mesure que l'on progresse vers les basses fréquences. Cette remarque concerne particulièrement les instruments dont l'énergie est concentrée à 90 % dans la fréquence fondamentale, comme c'est le cas pour les flûtes traversières. La flûte contrebasse (à l'octave inférieure de la flûte alto), instrument le plus grave de cette famille, qui descend au sol_1 (100 Hz)[6], nécessite un support et requiert le plus souvent une amplification. L'orgue est une exception à cette règle, puisque l'énergie n'est pas fournie par le musicien mais par une source d'air comprimé autonome, initialement produite par deux ou trois personnes actionnant des soufflets et aujourd'hui par un ventilateur électrique. Les plus gros consommateurs d'air sont les grands tuyaux à bouche (de type flûte) de 10 m de long (tuyaux de 32 pieds) dont la fréquence fondamentale descend au do_{-1} (16 Hz). Il faut bien reconnaître que des sons aussi graves jouent surtout un rôle de soutien de la polyphonie.

Le problème se pose différemment pour les instruments à anche dont le spectre est extrêmement riche en harmoniques. Le basson, le contrebasson, la clarinette contrebasse ou la voix humaine émettent des sons dont le fondamental et les premiers harmoniques sont très faibles, voire inexistants sur les notes les plus graves. Moins gourmands en énergie, ces instruments nous font entendre la fréquence fondamentale à travers les harmoniques aigus, c'est-à-dire sur la base de la périodicité détectée lors du codage temporel (chapitre 3 § 1.5.3). Bien que fonctionnant selon des principes différents, les instruments à cordes suivent les mêmes règles que les instruments à anche. La seule limite en tessiture est celle de la perception

Figure 5.8 L'octobasse construite par J. B. Vuillaume vers 1850. Longueur totale de l'instrument : environ 3,90 m. Un exemplaire est exposé au musée de la Musique à Paris.

Colomb C., 1878, *La musique*, Paris : Hachette, figure 117.

6. Il faut mettre à part les flûtes qui ne jouent que les partiels supérieurs comme la *fujara* slovaque dont le tuyau est de l'ordre de 1,70 m.

des basses fréquences par l'oreille. Citons aussi l'exemple de l'octobasse, instrument dont les dimensions sont telles que les notes sont produites par un système mécanique actionné au pied (figure 5.8).

Un son grave dont le fondamental est faible paraît grêle, quelquefois nasillard. Si l'on recherche des sons graves dont le fondamental est très intense, il faut fournir une énergie conséquente. Le goût actuel pour les musiques possédant beaucoup de basses s'est développé parce que les techniques de fabrication des enceintes acoustiques et le matériel d'amplification ont permis d'atteindre des niveaux élevés, capables d'emplir de grandes salles. Les auditeurs ont ainsi découvert des sensations nouvelles auxquelles ils sont maintenant attachés, au point que certains d'entre eux sont déçus par l'audition d'un grand orchestre qui, selon eux, « manque de basses » !

4. La dynamique des instruments de l'orchestre

4.1. La dynamique globale

La dynamique, ou écart entre les sons les plus faibles et les sons les plus forts qu'il est possible de jouer, se mesure ordinairement en dB (voir Glossaire). La figure 5.9 présente l'évolution de l'intensité des sons de trois instruments joués en gammes chromatiques détachées, du grave à l'aigu. La courbe rouge correspond au jeu *forte* et la courbe bleue au jeu *piano*. L'écart d'intensité entre ces deux nuances extrêmes présente des variations avec la tessiture qui diffèrent selon les instruments. La dynamique de la trompette est en moyenne de l'ordre de 15 dB. Elle atteint 20 dB dans le grave et sur certaines notes de l'aigu. Du grave à l'aigu, la dynamique se translate de 20 dB. La flûte traversière a une dynamique moyenne de l'ordre de 10 dB, très variable selon les notes, considérablement affectée par le vibrato. Les différences entre le grave et l'aigu atteignent 30 dB. La dynamique du violon, de l'ordre de 10 dB, reste stable sur toute la tessiture. Le violon, ainsi que les instruments à anche double, est le représentant le plus emblématique des instruments dont les variations de nuances, c'est-à-dire la dynamique musicale perçue, sont produites de diverses manières : par des changements de spectre, par des modifications de l'attaque du son et divers aspects du jeu comme le vibrato (voir page 217). Il n'est donc pas étonnant que la dynamique du violon mesurée en dB varie selon les interprètes (comparez les figures 5.9 et 5.10).

Les instruments qui ont la plus grande dynamique, ceux dont le contrôle des nuances d'intensité entre pour une part importante dans la technique de jeu, sont sans conteste les instruments à percussion, piano y compris. Le fait de choisir des percuteurs plus ou moins durs (bois, feutre) accroît la dynamique, tout en diversifiant les sonorités. Il est très difficile d'évaluer la sensation d'intensité d'un transitoire impulsionnel. Plus que l'écart total, en dB, c'est principalement le temps de montée (en dB/s) et le contenu en fréquences aiguës qui affectent la sensation d'intensité perçue.

Clark et Luce ont publié en 1965 une étude comparée de la dynamique de plusieurs instruments de l'orchestre. Les résultats, qui sont reproduits figure 5.10, présentent les allures caractéristiques que nous venons de décrire pour la flûte, la trompette et le violon. L'échelle en dB est relative mais permet de comparer les instruments entre eux, car ils ont été joués et mesurés dans les mêmes conditions acoustiques.

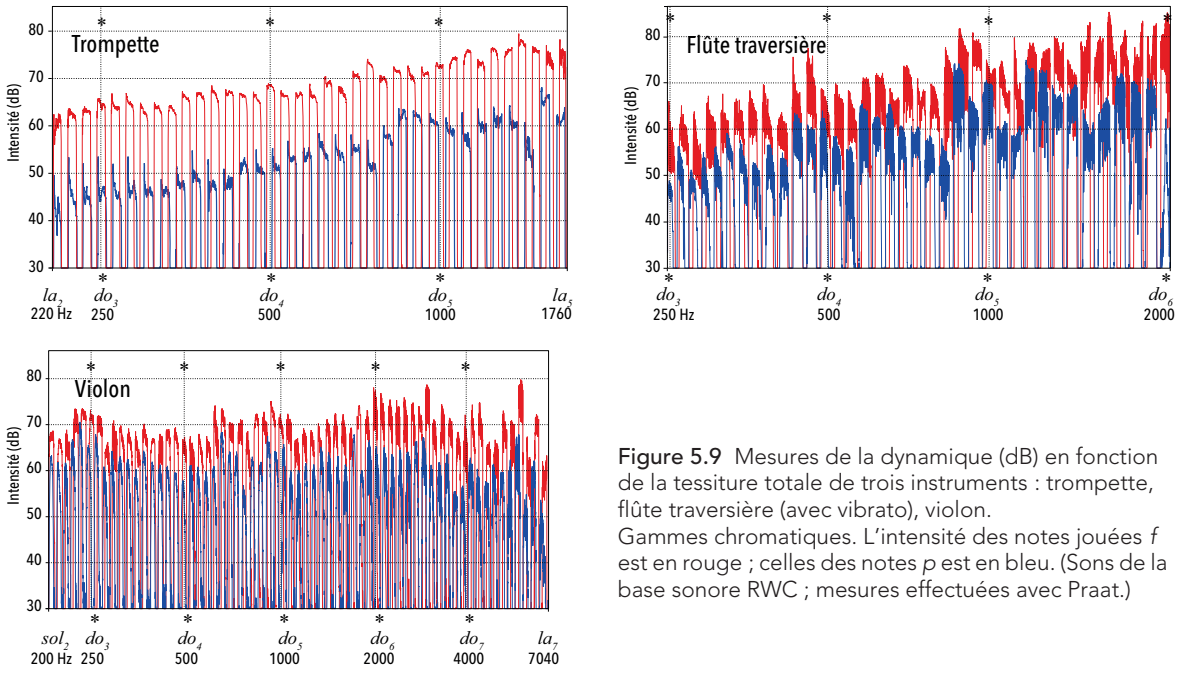


Figure 5.9 Mesures de la dynamique (dB) en fonction de la tessiture totale de trois instruments : trompette, flûte traversière (avec vibrato), violon.
Gammes chromatiques. L'intensité des notes jouées *f* est en rouge ; celles des notes *p* est en bleu. (Sons de la base sonore RWC ; mesures effectuées avec Praat.)

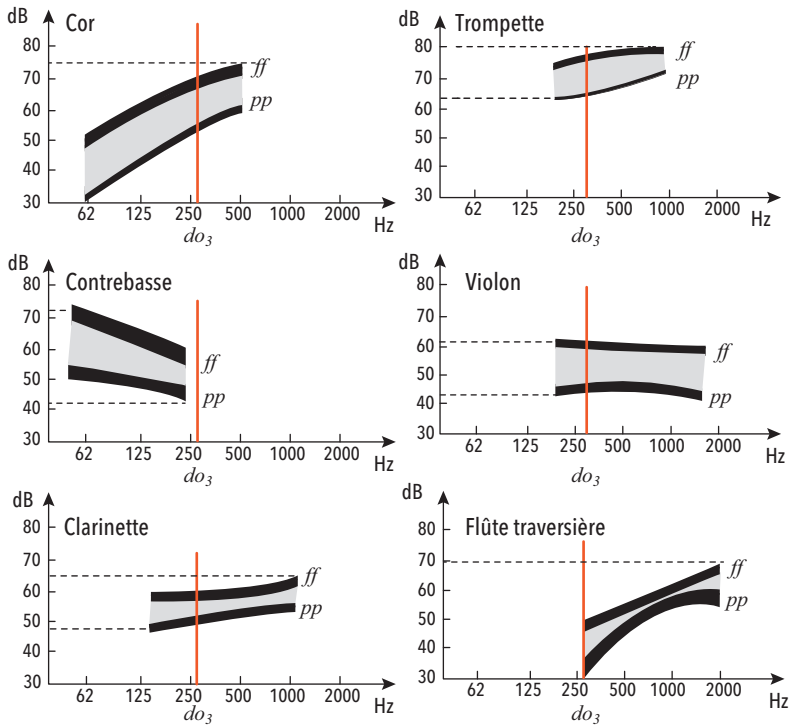


Figure 5.10 Dynamique de quelques instruments d'orchestre.

D'après Castellengo M., 1987, « Les sources acoustiques », in *Le livre des techniques du son* (Mercier éd.), Dunod (4e éd., 2010), figure 2.7.

La trompette se situe à la limite supérieure des intensités, entre 60 et 80 dB. La clarinette et le violon se positionnent dans une zone intermédiaire, respectivement 45-65 et 40-60 dB. Le cor et la flûte traversière partagent une dynamique qui varie fortement avec la tessiture. Entre les sons graves, faibles (30 à 50 dB) et les sons aigus intenses (55 à 75 dB), l'écart est de l'ordre de 25 dB. Notons que le diagramme de la contrebasse évolue en sens inverse, puisque les sons graves doivent nécessairement être produits plus fort pour être entendus.

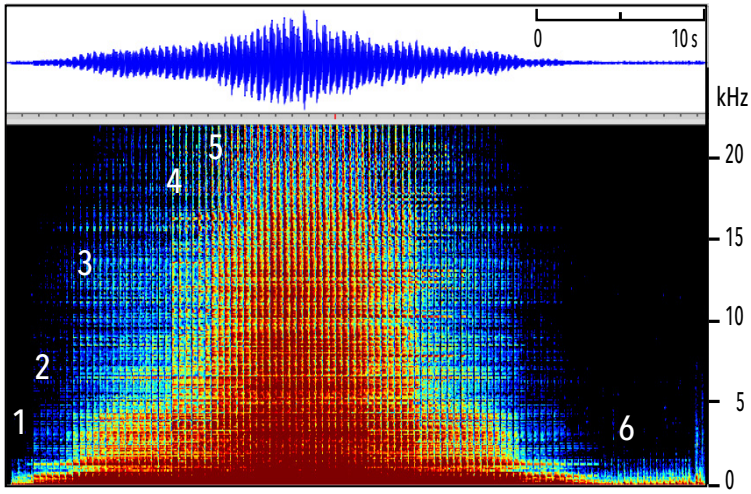
4.2. Les variations de dynamique par la registration

La mécanique de l'orgue et celle du clavecin ne permettent pas d'agir sur l'intensité des sons produits. Ces instruments servent une musique d'essence polyphonique dans laquelle les variations d'intensité résultent des modifications de la registration. D'une pièce à l'autre, le musicien modifie l'association des jeux ou des rangs de cordes. Sur un clavecin à deux claviers, l'augmentation de niveau sonore par accouplement, variable selon les instruments, est de l'ordre de 3 à 5 dB. Pourtant le remplissage spectral et le léger décalage des attaques entre les deux claviers suffisent à procurer une sensation d'intensité plus grande. Vers la fin du XVIIIe siècle, les facteurs ont ajouté, à l'imitation de l'orgue, des couvercles à volets pivotants, ou « jalousies », que l'on pouvait actionner en cours de jeu pour atténuer le son. D'autres ont réalisé d'ingénieux systèmes de genouillères qui permettaient l'actionnement des jeux en cours d'interprétation, sans quitter les claviers.

Son 5.7 (42'')

Effet de crescendo et decrescendo avec un clavecin à genouillères (détails page 228)

Voici un exemple de crescendo-decrescendo joué sur le clavecin Goujon du musée de la Musique, instrument à deux claviers et quatre registres de jeux (Son **5.7**, figure 5.11). Le musicien joue une suite d'accords répétés en commençant par le jeu de buffle[7], puis il introduit un à un les trois autres jeux (grand 8', petit 8', 4') et les retire pendant le decrescendo, pour finir avec seulement le son de la mécanique des claviers. L'entrée des jeux, qui est très progressive, et la distribution du nombre de notes dans les accords permettent de produire une gradation régulière de l'intensité. Du pianissimo au fortissimo, la dynamique totale est de 25 dB.



Crescendo : clavecin à deux claviers et à genouillères

Figure 5.11 Analyse spectrale d'un crescendo-decrescendo réalisé sur un clavecin à genouillères. (1) Bruit de fond ; (2) jeu de buffle ; (3) jeu de 8' ; (4) autre jeu de 8' ; (5) jeu de 4'. Ensuite retrait des jeux dans l'ordre inverse. En (6) bruit des touches des claviers. Échelle de fréquences 0 à 20 kHz.

7. Les cordes d'un jeu de buffle sont pincées par un morceau de cuir, ce qui produit un son plus doux. Comme à l'orgue, un jeu de 4 pieds (4') est à l'octave aiguë des jeux de base (8'). Le grand 8' du clavier inférieur est harmonisé différemment du petit 8' du clavier supérieur.

5. Jouer fort, jouer doux : le niveau sonore n'est pas seul en cause

5.1. Les transformations temporelles et spectrales

Son 5.8 (29″)

Violon nuance forte ; dix sons

Son 5.9 (46″)

Violon nuance piano ; neuf sons

Son 5.10 (8″)

Violon. Le son la_3 *joué successivement* f *et* p *(figure 5.12)*

Même dans le cadre d'un enregistrement technique, en l'occurrence celui d'une base de données, un musicien qui doit jouer avec des « intensités » différentes, fait appel à toutes les ressources que lui offre son instrument pour produire les nuances de jeu. Nous proposons tout d'abord d'écouter un arpège de *la mineur* joué au violon, avec la nuance *forte* (Son **5.8**) puis *piano* (Son **5.9**). Tous deux sont extraits d'une gamme chromatique en notes détachées, jouée par le même instrumentiste, dans les même conditions de prise de son. De nombreuses différences sont perceptibles à l'écoute et en particulier le fait que l'instrumentiste joue plus lentement la suite des sons en nuance *piano* qu'en nuance *forte.* L'analyse comparée de la figure 5.12 porte sur la note la_3 prélevée dans ces deux exemples sonores (Son **5.10**).

Indépendamment de la richesse spectrale du son *forte* mentionnée au § 3, d'importantes différences temporelles apparaissent. Le transitoire d'attaque du son fort, extrêmement bref, comporte des bruits de grincement (traits entre les harmoniques), alors que les harmoniques du son doux entrent progressivement, du grave à l'aigu, sans bruit. La différence globale d'intensité, qui n'est que de 8 dB dans la partie tenue (a), est ressentie principalement par la façon dont l'intensité globale varie dans le temps, principalement au début et à la fin du son.

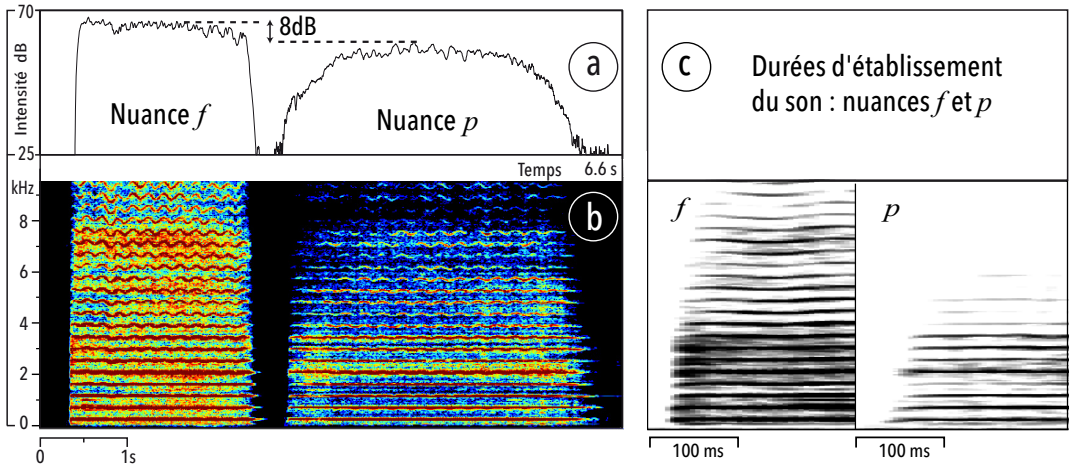


Figure 5.12 Analyse de la note la_3 (440 Hz) jouée selon deux nuances : *forte* et *piano* (même instrumentiste, même violon et mêmes conditions d'enregistrement). a) courbes d'intensité ; b) sonagramme. Comparez le spectre harmonique, le vibrato, le transitoire d'attaque, les bruits ; c) zoom spectrotemporel (250 ms) sur le transitoire d'attaque des sons *f* et *p*. (Mesures de l'intensité en dB avec le logiciel Praat)

Les oscillations du vibrato sont plus rapides (voir figure 5.13 page suivante) en nuance *forte* (5,5/s) qu'en nuance *piano* (4,8/s). L'excursion en fréquence est plus grande en *forte :* 2,2 % de la fréquence moyenne contre 1,5 % en *piano*[8]. La différence ressentie est aussi due à la façon graduelle dont croît le vibrato au début du son *piano* et pendant son extinction, alors qu'il s'impose immédiatement dès le début du son *forte*.

8. Mesures faites avec Praat. Un vibrato de 3 % couvre un intervalle d'un quart de ton de part et d'autre de la fréquence moyenne (ici, 440 Hz). Voir chapitre 6, § 4.3.

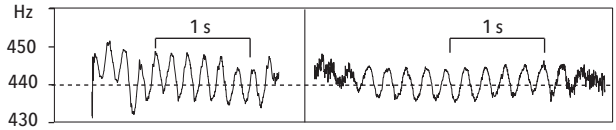


Figure 5.13 Tracé de la fréquence fondamentale (logiciel Praat) montrant les différences dans le vibrato en nuance *f* à gauche et en nuance *p* à droite.

L'ensemble de ces remarques confirme le fait que l'appréciation des nuances d'intensité concerne les plus fines variations de la forme globale spectrotemporelle des sons, ce qu'en acoustique on exprime par le fait que la sensation d'intensité, comme celle de hauteur et de timbre, est de nature multidimensionnelle.

5.2. L'estimation de l'intensité et la connaissance de la source

Un crescendo à niveau sonore fixe

Nous venons d'entendre qu'un violon jouant *pianissimo* a une sonorité caractéristique que l'on ne peut confondre avec celle du même instrument joué *forte,* ce qui explique qu'on peut ressentir les nuances musicales indépendamment de l'intensité globale.

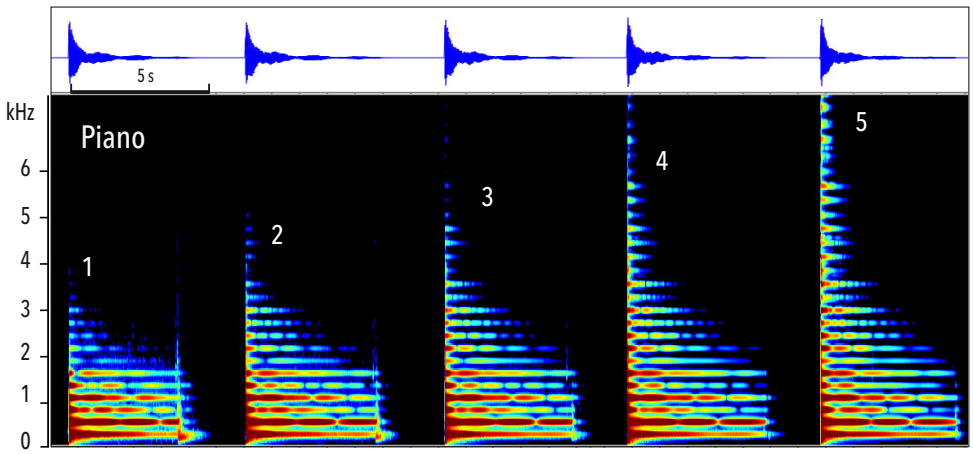


Figure 5.14 Analyse spectrographique de la note do_3 répétée *crescendo* puis égalisée en intensité. Notez les transformations de l'attaque, l'augmentation et l'enrichissement du spectre en partiels inharmoniques (Son 5.11a).

Piano : note répétée avec une frappe de force croissante

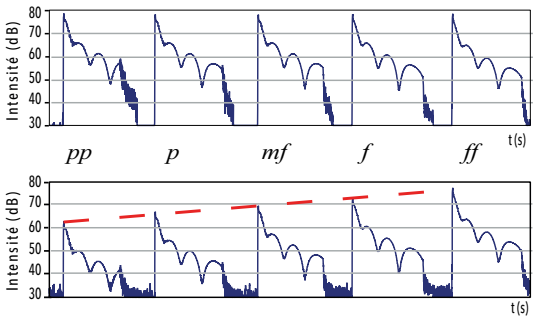
Son 5.11a (33'')

Niveau de chaque son normalisé

Voici un exemple (Son **5.11a**) enregistré sur un piano droit, instrument sur lequel le musicien ne peut régler que la vitesse d'enfoncement de la touche : c'est un crescendo sur la note do_3. L'écoute en est surprenante, car le signal enregistré a été modifié de sorte que les cinq répétitions de la note soient ajustées au même niveau sonore en dB, comme on peut le voir sur la courbe d'amplitude (figure 5.14).

Diffusés à la même intensité, les sons gardent l'information spectrale de la nuance initiale que nous avons mémorisée si bien que l'on entend, paradoxalement, un crescendo musical à intensité physique constante ! On voit sur le sonagramme de la figure 5.14 que le son s'enrichit en composantes aiguës lorsqu'il passe de 1 à 5, que l'attaque est de plus en plus nette et claire et que, d'une façon générale, les bruits de percussion augmentent.

La figure 5.15 montre les courbes de niveau sonore, en dB, calculées avec le logiciel Praat. La série supérieure correspond au Son **5.11a**, dans lequel les amplitudes ont été normalisées. Sur la courbe inférieure, qui est celle de l'enregistrement original du crescendo (Son **5.11b**), on lit que la différence d'intensité entre le son *pp* et le son *ff* est de 17 dB. Seule cette valeur est pertinente, car les unités affichées sur l'axe de gauche dépendent du matériel de lecture utilisé. Notez la régularité de l'accroissement d'intensité.



Son 5.11b (33″)

Captation sonore initiale

Figure 5.15 Intensité sonore, en dB, de la note de piano répétée en crescendo. Courbe supérieure : son normalisé en niveau sonore. Courbe inférieure : son original enregistré.

Pour parler doucement nous utilisons la voix chuchotée. Demandons à une locutrice de dire une phrase en voix chuchotée puis en voix normale, en s'ajustant au même niveau sonore à l'aide d'un décibelmètre (Son **5.12**). Curieusement la voix chuchotée paraît beaucoup plus forte que la voix normale alors que le niveau moyen des deux enregistrements est équivalent (figure 5.16).

Son 5.12 (7″)

Deux voix de même intensité globale (en dB)

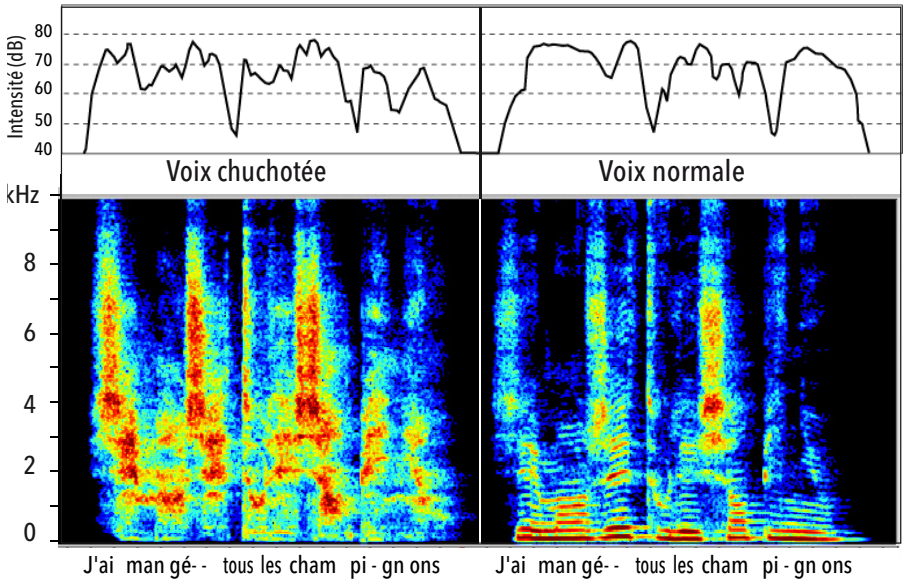


Figure 5.16 Une phrase parlée en voix chuchotée et en voix normale, à la même intensité moyenne. Intensité en dB (haut) et analyse spectrographique (bas).

Indépendamment du fait qu'une partie de l'énergie de la voix chuchotée affecte la zone sensible de l'oreille, divers signes (tonus de l'articulation, intonation) nous informent que la voix émise est forcée ou retenue, ce que nous traduisons aussi en sensation d'intensité.

Les indices acoustiques de l'intensité perçue sont spécifiques d'une catégorie de sons. Ils sont fondés sur notre connaissance des possibilités dynamiques propres à chaque type de source, ainsi que sur celle des variations spectrales et temporelles associées aux variations d'intensité.

5.3. La perception d'intensité en contexte temporel et le rôle des silences

Dans le cours d'une séquence musicale, nous sommes plus sensibles aux intervalles qu'aux valeurs moyennes : nous apprécions l'intensité d'un son par rapport à celui qui vient juste de se produire. Ainsi, lorsqu'un decrescendo succède à un crescendo la sensation d'intensité ne repasse pas nécessairement par les mêmes valeurs. On peut réaliser l'expérience suivante aisément au piano. Sur cet instrument, le contenu spectral et l'intensité étant étroitement liés, il est possible d'interpréter perceptivement la mesure physique de l'intensité. Jouons un accord étendu (aux deux mains) en le répétant 3 fois dans chaque nuance : *p, mf, f, mf, p*, puis mesurons le niveau sonore produit.

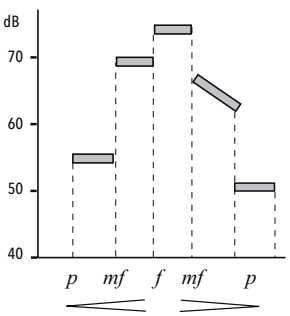


Figure 5.17 Accords répétés joués *crescendo* sur un piano.

Le résultat montre couramment la courbe présentée figure 5.17, c'est-à-dire que les niveaux *mezzo forte* et *piano* sont plus bas dans la partie descendante que dans la partie montante. Après un forte, l'oreille est un peu « éblouie » et il faut exagérer la variation d'intensité pour produire le même écart de nuance. Cet effet est implicitement pris en compte par les chefs d'orchestre.

Les silences, même brefs, modifient considérablement la perception d'intensité, sans doute en raison de la rapidité d'adaptation des muscles de l'oreille moyenne. On connaît la diversité des articulations consonantiques de la technique de jeu des instruments à vent, abondamment décrites dans les méthodes anciennes (Ganassi, 1535 ; Hotteterre, 1707), qui ont pour effet de moduler la durée des silences entre les « notes », tout en variant le contenu des transitoires d'attaque[9]. À l'orgue et au clavecin, instruments dits « inexpressifs », la durée des silences d'articulation entre les notes joue un rôle certain dans la sensation d'intensité des notes successives (voir Cohen, 2011). Nous avons pour témoignage les instruments mécaniques à cylindres du XVIIIe siècle et, plus particulièrement, les indications à l'adresse du « noteur » dont le travail est de reproduire la musique telle qu'elle est jouée. Or, on lit dans les ouvrages de l'époque : « toutes les notes (dans l'exécution) sont constituées de deux parties : une de tenue et une de silence. »[10]

9. L'arrêt du son produit par l'articulation d'une consonne plosive est plus long pour les « sourdes » (p, t, k) que pour les « sonores » (b, d, g). Voir chapitre 7, § 2.5.3.
10. J. Engramelle, 1775, *La tonotechnie ou l'art de noter les cylindres*, Paris ; Dom Bédos, 1776, *L'art du facteur d'orgue*, partie IV, chap.V.

L'effet de l'articulation par les silences est d'autant plus sensible au clavecin que l'étouffoir arrête le son de façon immédiate et précise, et que le transitoire qui succède au silence est très incisif. L'exemple sonore **5.13a** donne à entendre trois interprétations d'un thème de fugue. Dans la troisième interprétation, le contrôle fin de la durée des silences procure un phrasé dynamique simulant des accents d'intensité (figure 5.18 et Son **5.13b**).

Son 5.13a (22″)

Clavecin : trois exemples d'articulation son-silence

Son 5.13b (4″)

Fragment de la 3e interprétation

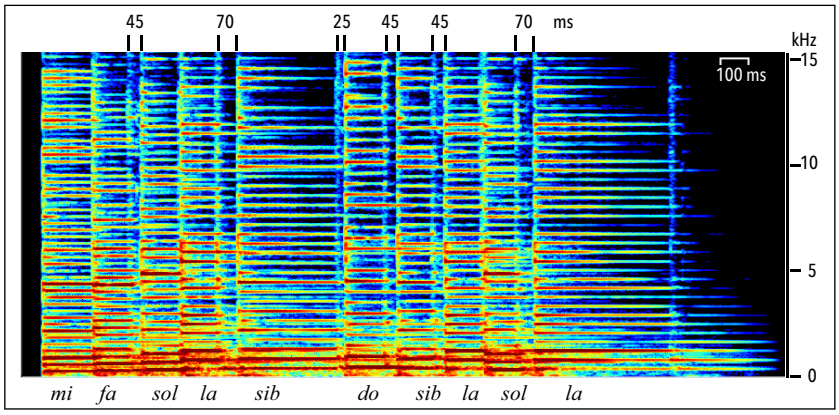


Figure 5.18 Analyse des deux groupes de doubles croches de la 3e interprétation. Remarquez la richesse spectrale du clavecin et la netteté du début et de la fin du son. Les silences varient de 25 à 70 ms, la durée la plus longue précédant la note la plus accentuée.

6. Les nuances d'intensité en musique

Les indications de nuance d'intensité sont extrêmement rares dans les partitions anciennes. Il s'agit essentiellement de l'opposition fort/doux mise en œuvre par l'instrumentation. Pour obtenir des variations d'intensité, on emploie alternativement les « hauts » instruments : trompette, hautbois[11] et percussion, et les « bas » instruments : flûtes et violons. Le même type d'opposition se produit au sein d'un groupe d'instruments par division des effectifs, dans le but de contraster deux parties d'une phrase musicale (voir figure 5.19 page suivante).

Cette pratique d'opposition, qui est très raffinée au XVIIIe siècle, implique que la musique se tienne dans une nuance d'ensemble donnée pendant une durée appréciable : nous dirons qu'il s'agit de nuances à moyen ou long terme. Ce goût se perpétue dans la formule du concerto pour soliste et ensemble orchestral.

Toute autre est l'appréciation des variations d'intensité à très court terme pour lesquelles nous avons une extrême sensibilité : accents, phrasés, gonflements sur une note, variations expressives que les musiciens pratiquent depuis toujours. De tels phrasés se sont considérablement développés au XIXe siècle, provoquant le déclin des instruments ne permettant pas de varier l'intensité de chaque note (clavecin, flûte à bec[12]), au profit d'autres, plus adaptés au nouveau goût (piano, flûte traversière), et entraînant la transformation d'un instrument pourtant très représentatif de l'ancien style, l'orgue, qui s'est alors doté de « boîtes expressives » et de systèmes mécaniques permettant d'introduire des jeux de façon insensible, afin de varier l'intensité au cours d'une même phrase.

11. Ce que révèle l'étymologie, l'orthographe ancienne étant : « haut-bois ».
12. Le fait de souffler plus fort à la flûte à bec pour produire un son plus intense entraîne automatiquement une montée de la note. Cet effet est évité dans le jeu classique, au profit des silences d'articulations et des ornements qui permettent de réaliser des accents dynamiques.

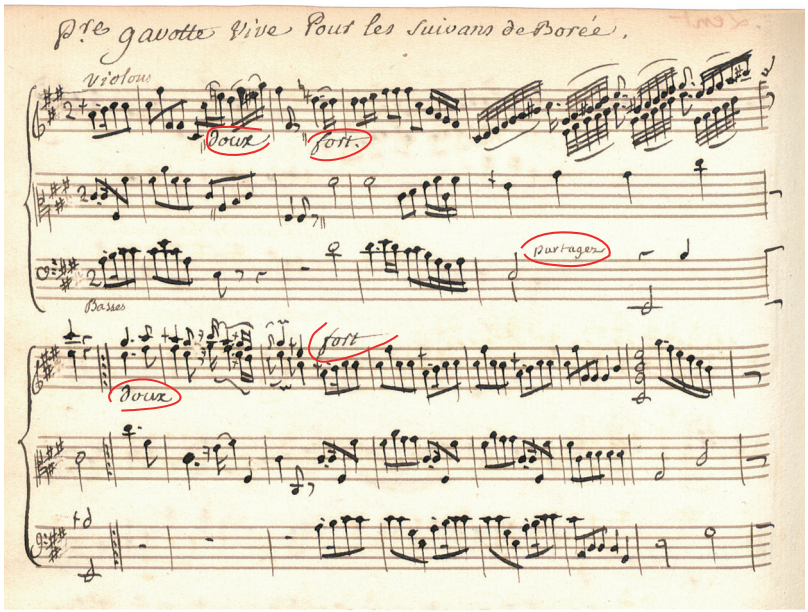


Figure 5.19 La dynamique et les nuances de jeu. Partition du XVIIIe siècle : Rameau, *Les Boréades*, 1764.

Rameau J.-Ph., manuscrit autographe, acte II, scène VI (Bnf).

Peu à peu, la notation des nuances s'est développée et complexifiée de façon indépendante pour chaque partie et avec une précision d'indication qui peut laisser perplexe (voir figure 5.20, la partition de Ligeti, 1970).

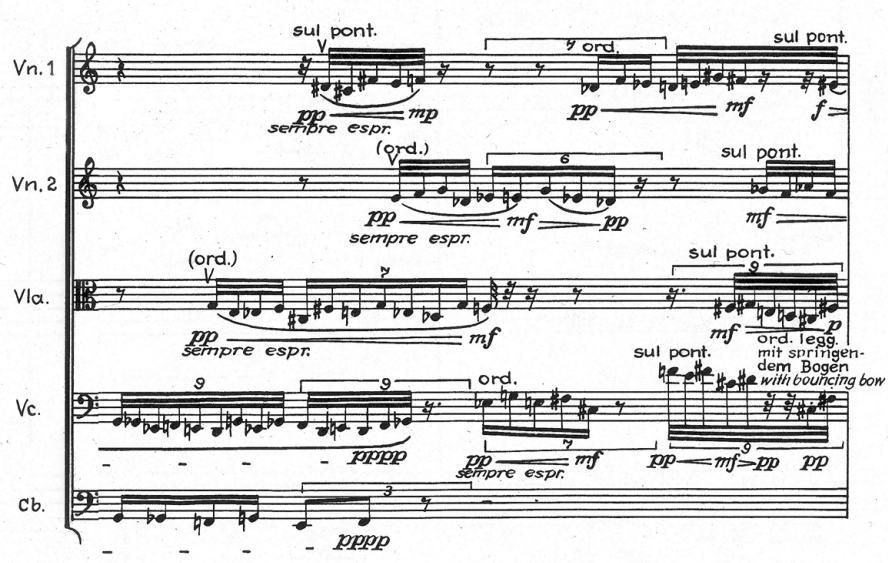


Figure 5.20 La dynamique et les nuances de jeu. Partition du XXe siècle : Ligeti, *Kammer-konzert für 13 instrumentalisten* ; 1970.

Ligeti G., 1974, SCHOTT Music, Mayence (Allemagne), page 27, fragment de la mesure 56.

7. L'émergence par contraste de forme

7.1. Les signaux d'avertissement

En présence d'un environnement sonore intense (bruit de foule, circulation), le premier réflexe pour se faire entendre consiste à augmenter le volume du son. Les ouvrages d'acoustique stipulent que pour passer au-dessus d'un bruit, il faut qu'un signal sonore ait un niveau de 6 dB supérieur à celui du bruit.

Quelques Parisiens se souviennent sans doute des premiers essais d'introduction d'un signal sonore de fermeture des portes sur la ligne 1 du métro. Pendant quelque temps, les voyageurs ont dû subir un son pur, stable, fatiguant, dont le niveau sonore avait été calculé pour être deux fois plus élevé (+ 6 dB) que le niveau moyen du bruit de fond (de l'ordre de 80 dB). Ce signal a heureusement été remplacé par un signal moins intense, riche en harmoniques et animé de battements dont la forme spectrotemporelle, immédiatement reconnaissable, émerge parfaitement sur le bruit de fond ambiant sans agresser l'oreille (figure 5.21 et Son **5.14**).

Son 5.14 (27'')

Métro : émergence du signal sonore de fermeture des portes (1980)

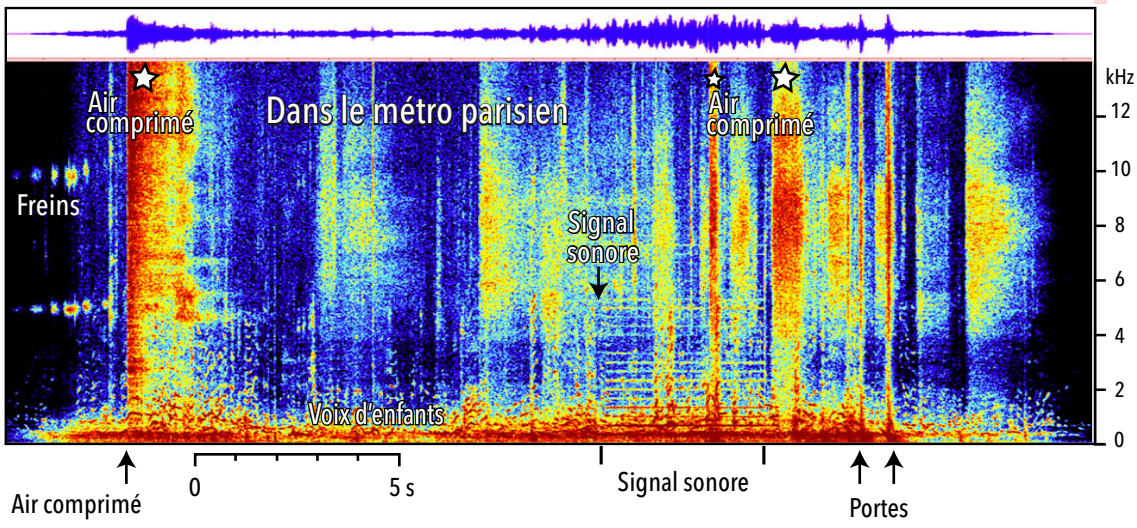


Figure 5.21 Le signal sonore du métro sur le bruit de fond. Par sa structure harmonique (raies parallèles), un signal sonore musical émerge sur un bruit de fond intense et complexe couvrant tout le spectre sonore.

C'est en effet essentiellement par les variations *temporelles* de spectre et de fréquence, celles qui constituent la « prégnance acoustique » d'une forme sonore par rapport au bruit de fond, que la plupart des signaux d'avertissement – produits traditionnellement par des sources mécaniques, de petites dimensions et dont la fréquence cible la zone de plus grande sensibilité de l'oreille –, émergent avec efficacité.

E. Leipp a montré, dans une étude pionnière (1965), la remarquable adaptation des signaux produits par les sifflets d'avertissement dont la forme spectrotemporelle du signal codé variait selon la diversité des environnements sonores professionnels : marine, rail, circulation automobile (figure 5.22).

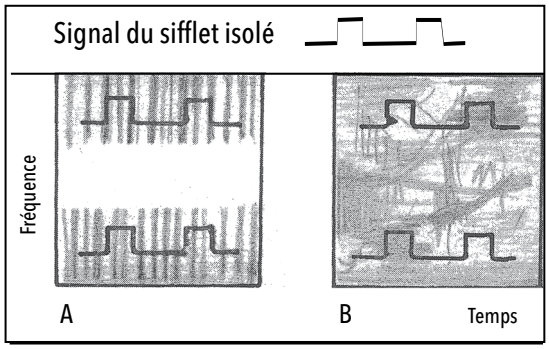


Figure 5.22 Formes et fond. Emergence acoustique d'un sifflet à deux notes sur deux bruits de fond différents : « hachuré » (sons percussifs) et « continu erratique » simulant la circulation.
Document Emile Leipp, archives LAM.

Les signaux d'alerte modernes, amplifiés, génèrent tous des sons très riches en harmoniques. Il suffit que quelques composantes du spectre émergent pour que le « message », codé par une forme mélodicorythmique, reste perceptible (voir le chapitre 6, § 2.3).

7.2. La notion d'émergence en musique

Son 5.15 (24'')
Cabrette d'Auvergne et grelots

Un problème similaire est celui des musiques de plein air dont la mélodie et le rythme doivent émerger par rapport au fond sonore d'une foule et pour lesquelles les sonorités des instruments ont été adaptées. Deux stratégies existent : soit concentrer l'énergie dans la zone de plus grande sensibilité de l'oreille par l'emploi d'une petite flûte (galoubet provençal, txistu basque), soit utiliser une anche double dont la richesse harmonique permet de passer au-dessus de la limite supérieure en fréquence du bruit de la foule (bombardes et binious bretons ; cabrette auvergnate). L'émergence du rythme est assurée de diverses façons (tambours, grelots). Écoutez l'association cabrette et grelots (Son **5.15**) dont l'analyse est donnée figure 5.24.

Figure 5.23 Joueur de cabrette (Jean Rascalou).
Wikipedia ; vers 1910, auteur inconnu.

De façon similaire, la voix chantée se détache aisément de l'ensemble des instruments par la présence d'un vibrato plus ample que celui des violons ou des flûtes. On peut voir (figure 5.24) et entendre (Son **5.16)** l'émergence de la voix d'une soprano, sans microphone d'appoint, même dans le *tutti forte* qui précède la phrase soliste.

Son 5.16 (20'')

Émergence de la voix chantée sur un grand orchestre

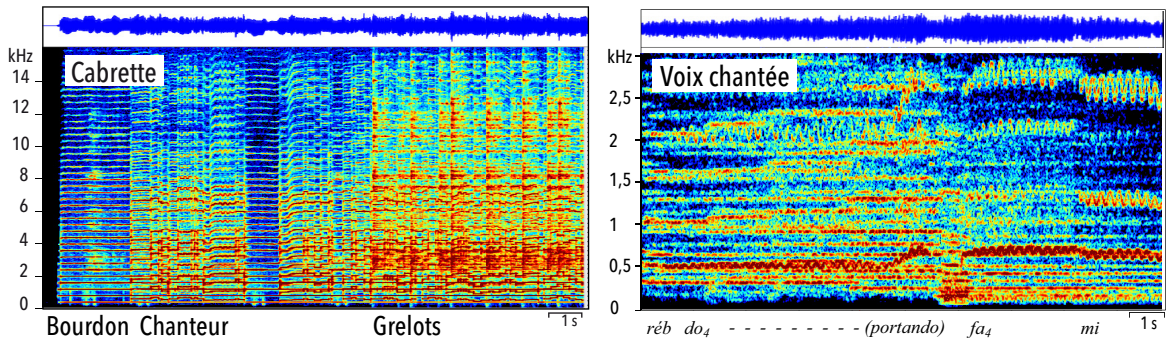


Figure 5.24 À gauche, jeu de cabrette auvergnate (Son 5.15). L'apport des grelots en termes de niveau sonore ne dépasse pas 3 dB, mais leur émergence perceptive, bien visible sur l'analyse spectrotemporelle, est due à trois caractéristiques : impacts précis qui tranchent sur la mélodie du tuyau chanteur ; large bande spectrale (jusqu'à 20 kHz), cavités de résonance accordées et centrées sur 3 kHz. À droite : une chanteuse soliste émerge sur un ensemble orchestral par l'allure caractéristique de son vibrato, même lorsque l'orchestre atteint le *ff* (Son 5.16). (Enregistrement en *live* avec un seul micro)

À l'orchestre le triangle émerge de la masse instrumentale, bien qu'il rayonne peu d'énergie. Le choc dû à la frappe et le contenu spectral aigu de l'instrument – situé dans la zone de 2 à 4 kHz) – constituent une forme spectrotemporelle qui sort de la zone de fréquence des principaux instruments de l'orchestre.

Les « distorsions » spectrales générées par des dispositifs placés intentionnellement sur certains instruments pour produire un enrichissement spectral offrent une solution élégante pour émerger d'un ensemble instrumental. Citons le « timbre » du tambourin provençal (Castellengo et Fabre, 1966) ; le chevalet instable (ou « chien ») de la vièle à roue et de la trompette marine ; l'obstacle ponctuel placé sur le parcours de l'impulsion d'une corde pincée : le *jouari* de la tampoura indienne (Bertrand, 1992) et les « harpions » de la harpe Renaissance (Musson-Gonnaud, 2013) ; le « mirliton » des flûtes asiatiques.

Le contrôle temporel de l'action de ces dispositifs en cours de jeu offre la possibilité de produire des rythmes spectraux tout à fait efficaces, en particulier lorsqu'il s'agit d'émerger d'un ensemble instrumental ou sur le bruit d'une foule.

Nous avons rassemblé sur la figure 5.25, page suivante, quelques exemples d'émergence spectrale adaptés à différentes musiques : glissando spectral de la tampoura, roulement du tambourin de Provence, saturation spectrale de la vièle à roue et grelots cristallins de la cabrette (Son **5.17**).

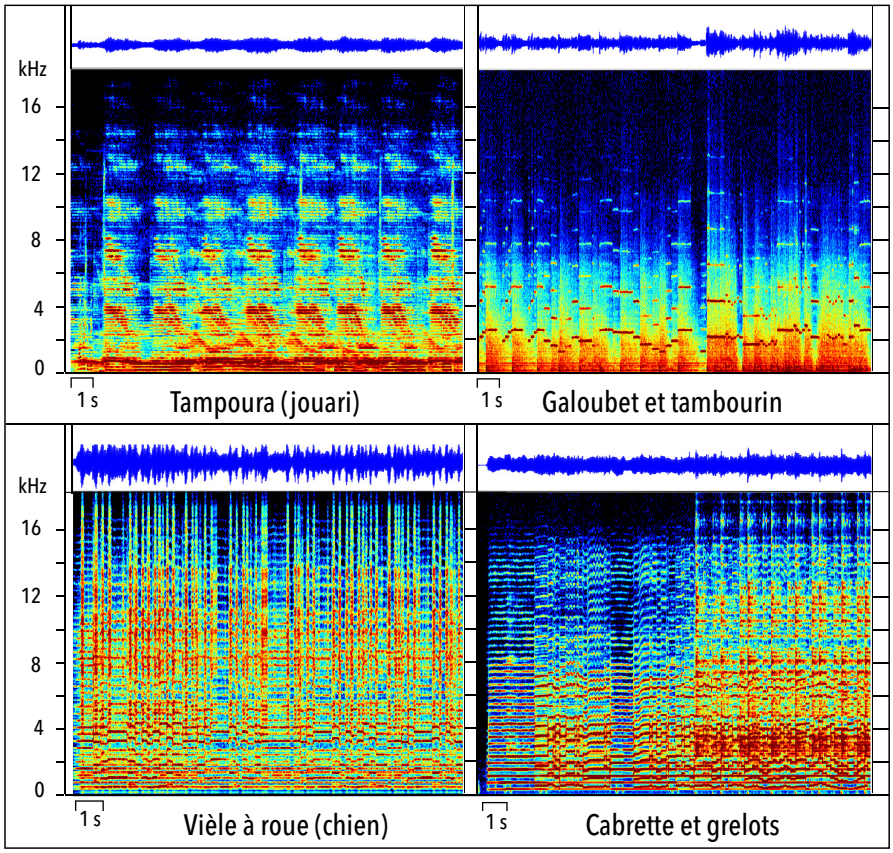


Figure 5.25 Exemples d'associations mélodie et rythme produites par un seul musicien, avec peu d'énergie et une grande efficacité perceptive. Analyses de 0 à 18 kHz.

Son 5.17 (1'16'')

Exemples musicaux de la figure 5.25

Mentionnons pour finir les rôles de l'attention et de la focalisation perceptive qui permettent à un auditeur d'accentuer l'émergence des sons qu'il désire particulièrement entendre. Tout musicien peut « suivre » une source musicale dans un *tutti* d'orchestre ou dans un ensemble choral polyphonique en développant une capacité qui mobilise tout le système nerveux central, capacité encore difficile à expliquer aujourd'hui.

8. Conclusion

L'intensité joue un rôle majeur dans la perception des sons. Porteuse d'indices associés aux émotions fondamentales de peur, d'agressivité ou de douceur, la sensibilité aux changements d'intensité d'un son est ressentie par tous et repose en premier lieu sur la relation entre l'augmentation d'énergie à la source et l'accroissement de l'amplitude vibratoire : pour jouer plus fort il faut donner plus d'énergie. Celle-ci étant limitée, une première parade consiste à la placer dans la zone des 3 000 Hz, là où l'oreille est la plus sensible, ce qui produit des sons aigus souvent agressifs. En musique où les variations de l'amplitude sonore affectent des systèmes mécaniques non linéaires et très élaborés, nous avons vu que les changements d'intensité du jeu instrumental se traduisent par de complexes modifications du spectre

– comme l'accroissement du nombre des harmoniques dans le crescendo – mais aussi, et plus particulièrement les chanteurs, en ciblant électivement les composantes situées autour de 3 000 Hz.

La complexité des effets dus aux variations d'énergie à la source explique qu'il n'existe pas de notation précise pour la sensation d'intensité sonore et que le décibelmètre, sensible à la seule amplitude de l'onde sonore, n'en rende compte que de façon très imparfaite. En pratique les musiciens s'entendent globalement sur cinq très grands échelons qui vont du pianissimo au fortissimo. Mais c'est sans compter sur l'effet des faibles changements d'intensité à court terme, ressentis avec une grande finesse de résolution, puisqu'un écart de 1dB (à 1000 Hz pour un niveau moyen de 50 dB) est déjà bien perçu par l'ensemble des auditeurs. À cela s'ajoutent les subtilités du jeu expressif des interprètes qui par le geste d'attaque du son, sa conduite temporelle et par la modulation des silences d'articulation affectent aussi le ressenti d'intensité.

Le fait que certains sons émergent sur un ensemble complexe dépend moins de leur intensité que de l'originalité de leur forme spectrale et de ses variations temporelles par rapport au fond sonore : plus cette forme est prégnante et plus il est aisé de suivre un élément spécifique dans un *tutti*. De ce point de vue les signaux d'avertissement et les musiques traditionnelles représentent des solutions élégantes et éprouvées, qui mettent en jeu peu d'énergie du fait de leur remarquable adaptation aux propriétés de l'oreille et aux stratégies du traitement cognitif.

9. Les sons du chapitre 5

9.1. Intensité perçue et spectre

Son 5.1 – Phrase chantée par un baryton : « un sentiment de crainte ». Son 5.1a : voix naturelle. Son 5.1b : voix dans laquelle le formant du chanteur a été complètement filtré entre 2200 Hz et 3300 Hz. Baryton, 1969. [Archives LAM]

Son 5.2 – Une voix répond au téléphone : enregistrement à la sortie du récepteur, M. C., 1986. [Archives LAM]

Son 5.3 – La même voix enregistrée simultanément à la source. La comparaison avec l'exemple précédent met en évidence les transformations spectrales opérées lors de la transmission : la qualité vocale est profondément modifiée. Nous sommes y tellement habitués que nous n'y faisons plus attention, puisque la parole est intelligible et que les intonations sont conservées. M. C. [Archives LAM]

Son 5.4 – Avant l'écoute du Son 5.4, assurez-vous que le niveau sonore est plutôt faible, et ne le modifiez pas pour le deuxième son. Cet exemple fait entendre deux sons ayant la même fréquence fondamentale, $f = 500$ Hz, et le même niveau sonore en dB. Le Son 5.4a est un son pur. Le Son 5.4b est un son complexe ayant cinq harmoniques. Écoutez une deuxième fois ces exemples en augmentant le niveau sonore. Vous pourrez constater que la différence d'intensité perçue est plus notable lorsque le son est fort. [M. C.]

9.2. Crescendo musical et changement spectral

Son 5.5 – Une note jouée à la trompette, sib_3, est répétée avec augmentation progressive de l'intensité. Louis Roquin, 1977. [Archives LAM]

Son 5.6 – Son filé joué au saxophone alto : sib_2 = 230 Hz. Le crescendo-decrescendo dure 21 secondes. Claude Delangle, 1989. [Archives LAM]

Son 5.7 – Séquence jouée sur un clavecin muni de genouillères permettant d'actionner les registres et l'accouplement des claviers pendant le jeu. Le musicien joue des accords des deux mains, dont il adapte le contenu en fonction du crescendo qu'il veut produire. Il commence sur le jeu de buffle, puis introduit le grand 8', met l'accouplement, ce qui fait intervenir le jeu de 8' du 2e clavier, puis le jeu de 4'. Le decrescendo se fait dans l'ordre inverse et l'exemple se termine avec le seul bruit des touches des claviers. Exemple enregistré au Musée instrumental du CNSM ; D. Füller, 1980. [Archives LAM]

Son 5.8 – Violon joué en nuance *forte*. Arpège de *la* mineur sur trois octaves (la_2-la_5). Sons extraits de la gamme chromatique 151VNNOF : base de données RWC. [M. C.]

Son 5.9 – Violon joué en nuance *piano*. Arpège de *la* mineur sur trois octaves (la_2-mi_5). Même violon et même interprète que pour l'exemple précédent joué *forte*. Sons extraits de la gamme chromatique 151VNNOP : base de données RWC. [M. C.]

Son 5.10 – Violon : écoute comparée de la même note, la_3, jouée *f* puis *p*. Sons extraits des arpèges des sons 5.8 et 5.9. [M. C.]

Son 5.11 – Son 5.11a : une note de piano, do_3, a été jouée selon cinq nuances d'intensité : *pp, p, mf, f, ff*. Ensuite les cinq sons enregistrés ont été ramenés à la même intensité physique en dB. Malgré cette opération, les modifications de la forme spectrotemporelle du son – qui sont considérables – continuent de porter l'information d'un crescendo sonore. Son 5.11b : la séquence originale enregistrée au piano, avant égalisation ; M. C., 2007. [Archives LAM]

Son 5.12 – Voix parlée. Successivement : voix chuchotée et voix normale. Les deux exemples vous paraissent-ils à la même intensité ? Nathalie Henrich, 2001. [M. C.]

Son 5.13 – Phrase musicale jouée au clavecin. Son 5.13a : interprétation avec trois articulations de silences différentes. Successivement : jeu entièrement legato ; jeu alternant staccato et legato ; jeu avec articulations différenciées. Son 5.13b : extrait sonore correspondant à l'analyse sonagraphique de la figure 5.18. Thème de la *Fugue n° 11 en* fa *majeur* ; *Clavier bien tempéré*, Vol. 2, J.-S. Bach. L'instrument est au diapason 415 Hz. A. Geoffroy-Dechaume, 1980. [M. C.]

Son 5.14 – Courte séquence enregistrée dans un wagon de métro où se trouve un groupe d'enfants. Bruits divers : crissement des roues ; sifflements d'air comprimé ; claquements d'ouverture et de fermeture des portes. Le signal sonore du départ du train, qui est un son riche en harmoniques, émerge clairement du bruit ambiant. Paris, ligne 7, 1980. [M. C.]

Son 5.15 – Air de cabrette : musique traditionnelle d'Auvergne. Le musicien porte deux bracelets de grelots aux pieds avec lesquels il marque le rythme peu après le début du jeu. Jean Bona, 1981. [Archives LAM]

Son 5.16 – Voix chantée et orchestre. Enregistrement *live*, 1955 ; Maria Callas, Violetta, *La Traviata*, Acte 2, *Dammi tu forza*. [G. Bloch.]

Son 5.17 – Les quatre extraits de musiques analysées figure 5.25. Successivement : tampoura de l'Inde du Nord (1967, réunion du GAM n° 31) ; galoubet et tambourin (Marius Fabre 1966, réunion du GAM n° 23) ; vièle à roue ; cabrette et grelots (J. Bona, 1981). [Archives LAM]

CHAPITRE 6

PERCEPTION DES QUALITÉS SONORES : LA HAUTEUR DES SONS ISOLÉS

1. Préliminaires

1.1. Le paramètre « hauteur » dans la musique occidentale

Dans la culture occidentale, la hauteur a donné lieu au développement de musiques mélodiques, contrapuntiques et harmoniques d'une grande richesse. Cette situation est l'aboutissement d'une lente élaboration multiséculaire affectant à la fois les instruments, la notation de la musique et la pédagogie musicale.

L'histoire de la musique montre une tendance constante de l'utilisation de sons harmoniques, accordables, permettant une perception bien définie et précise de la hauteur. Celle des transformations de la facture des instruments de musique ainsi que l'inventaire des techniques de jeu confirment la recherche incessante d'innovations pour qu'il soit possible de jouer aisément des notes justes, repérables, quel qu'en soit l'instrument[1].

L'écriture musicale privilégie la représentation graphique de la hauteur aux dépens de celles d'autres variations de qualité des sons comme l'intensité ou le timbre. La notation graphique du mouvement mélodique, puis celle des degrés de l'échelle musicale repérés par rapport à un système de lignes horizontales, a joué un rôle de renforcement et de focalisation sur la dimension de la hauteur, et a scellé l'association « hauteur et verticalité » stabilisée au début du XIXe siècle avec l'adoption générale de la portée au détriment des tablatures, système qui représente la position des doigts sur l'instrument, ce qui permet de jouer facilement des instruments de dimensions différentes.

1. Citons les perfectionnements de la fabrication des « cordes harmoniques » et les importantes recherches développées au XIXe siècle pour améliorer la justesse des instruments à trous latéraux (flûtes, hautbois et clarinettes).

Parallèlement au développement d'une notation musicale de plus en plus précise – y compris celle des ornements mélodiques – s'élabore une pédagogie musicale qui conduit à privilégier la représentation mentale de *notes,* qu'il s'agisse de sons plus ou moins riches en harmoniques, de sons stables ou vibrés. Le musicien débutant doit chanter une note, lue sur une portée, en la nommant, et en y associant éventuellement un son instrumental de même hauteur. Ainsi s'élabore, selon l'expression de Robert Francès, « l'abstraction notale » ou capacité à entendre sélectivement la hauteur à travers la diversité des réalisations instrumentales. La notation écrite des sons musicaux qui, par ailleurs, a joué un rôle fondamental dans les développement des compositions occidentales, renforce cette schématisation perceptive.

Une telle focalisation culturelle permet, certes, de former des experts de haut niveau, cependant elle contribue à renforcer une audition sélective et réductrice, dans le sens où l'effort accompli pour « ignorer » la diversité des productions musicales d'une note aboutit à une sorte de surdité esthétique vis-à-vis des ornementations mélodiques développées dans les musiques d'autres cultures. Elle conduit aussi à écarter du champ instrumental classique des sons incongrus comme les sons multiphoniques, les sons d'instruments à percussion et, d'une façon générale, les sons de hauteur ambiguë que l'on trouve dans bon nombre de musiques traditionnelles. Leur réhabilitation dans la musique contemporaine, la confrontation aux sons des musiques du monde traditionnelles[2] et la création de sons « inouïs » par synthèse numérique conduisent à proposer un cadre plus général pour rendre compte de la perception de la hauteur.

La sensation de « hauteur » ne serait donc pas une donnée immédiate de la perception mais une compétence acquise. Oublions un instant nos connaissances, oublions la mesure de la fréquence pour retrouver la fraîcheur d'une écoute, sinon naïve, du moins curieuse, à l'audition des exemples suivants.

1.2. La perception de la hauteur sonore : trois expériences introductives

Les premières expériences de psychoacoustique ont montré que la sensation de hauteur est liée à la fréquence. Ce fait a été établi pour les sons purs et plus généralement pour les sons parfaitement périodiques. Si, munis de ce postulat, nous nous fondons sur la seule mesure de la fréquence fondamentale pour rendre compte de la hauteur des sons réels de la musique, nous obtenons quelquefois des résultats surprenants, voire des paradoxes incompréhensibles. Pourquoi certains sons semblent-ils baisser lorsqu'on les transpose vers l'aigu ? Que se passe-t-il lorsqu'on supprime la fréquence fondamentale ? Comment expliquer qu'un seul son évoque deux ou trois hauteurs différentes ? Enfin, pourquoi entendons-nous objectivement une hauteur bien présente alors qu'aucune source ne vibre à la fréquence de la note entendue ?

En guise d'introduction, nous invitons le lecteur à se prêter aux trois petites expériences d'écoute suivantes.

2. À ne pas confondre avec « world music », qui est plutôt une transposition du système harmonique occidental.

1.2.1. Expérience 1 : mélodies de sons bruités

Le premier exemple (Son **6.1**) présente quatre sons extraits de l'environnement quotidien, immédiatement reconnaissables à la première écoute : un marteau piqueur, de l'eau qui s'écoule, un bruit de machine (aspirateur), une goutte d'eau. L'analyse spectrale en est donnée ci-contre (figure 6.1).

Son 6.1 (6")

Quatre sons de l'environnement

En opérant deux transpositions en fréquence, nous obtenons l'exemple suivant (Son **6.2**) dans lequel chacun des sons précédents est répété trois fois (figure 6.2). Nous prenons soudain conscience du basculement qui s'opère : l'attention se porte maintenant sur les changements de qualité qu'a produits la transposition.

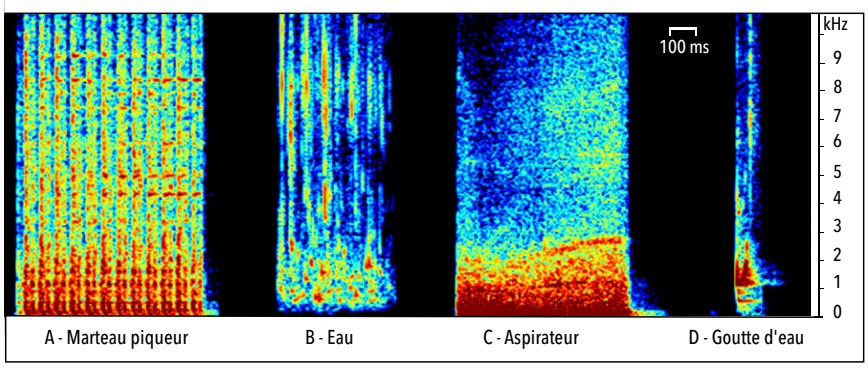


Figure 6.1 Représentation spectrale de 4 sons de l'environnement.

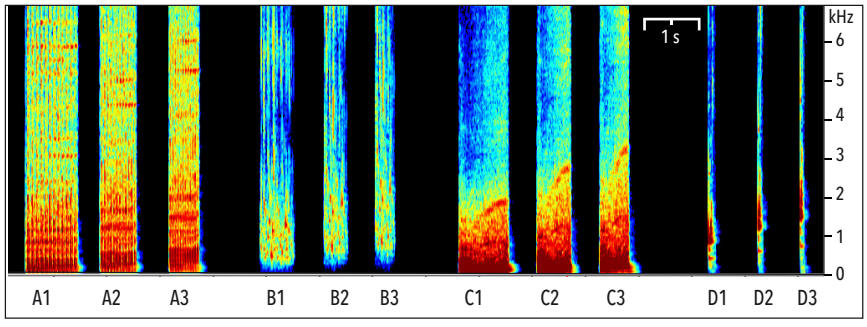


Figure 6.2 Transposition en fréquence. de chacun des quatre sons de la figure 6.1.

De la comparaison de sons similaires par leur origine naît soudain une relation que nous cherchons à décrire et à évaluer : le deuxième son de chaque groupe est « plus aigu, plus clair, plus haut, plus fort, etc. ». Nous sommes passés d'une écoute causale (Son **6.1**) à une écoute qualitative (Son **6.2**), de l'identification de la source à l'écoute musicale. En créant des suites de sons ayant la même origine causale, mais dont les qualités diffèrent, nous concevons déjà quatre « instruments de musique » jouant chacun trois sons. La sensation induite par la transposition spectrale est complexe, difficile à préciser pour les sons A, B, et C dont le spectre est étendu. Pour certains auditeurs, le son marteau piqueur (ou celui du ruissellement d'eau) ne suit pas les mêmes variations que les autres sons, alors que pour tous les auditeurs la goutte d'eau donne une sensation de montée vers l'aigu, nette et

Son 6.2 (16")

Mêmes sons répétés et transposés

appréciable : elle joue une mélodie. En choisissant des sons analogues nous allons maintenant tenter de comparer les intervalles de variation.

1.2.2. Expérience 2 : intervalles de hauteur entre deux sons

L'expérience porte maintenant sur trois sons dont la hauteur est plus facile à appréhender : le son produit par le débouchage d'une bouteille de vin, celui d'une bouteille de champagne et un son de flûte[3]. Comme précédemment, le deuxième son de chaque paire est obtenu par transposition du premier vers l'aigu.

Son 6.3 (8")

Trois sons « à note » transposés : estimez l'intervalle.

À l'écoute de chacun des couples du Son **6.3**, tous les auditeurs perçoivent un intervalle ascendant similaire. Toutefois, les auditeurs musiciens remarquent des différences entre ces intervalles, et certains peuvent même les nommer très précisément. Voici les résultats mentionnés par le plus grand nombre d'entre eux :

- intervalle des sons 1 et 2 : une quarte (5 demi-tons) ;
- intervalle des sons 3 et 4 : une quinte (7 demi-tons) ;
- intervalle des sons 5 et 6 : un triton (6 demi-tons).

Ces résultats sont très intrigants, car la transposition appliquée aux fréquences est strictement la même pour les trois sons. Comment se fait-il que nous entendions trois intervalles différents ? Comment se fait-il que les intervalles perçus changent selon les auditeurs ?

Selon nous, la raison principale tient à la structure acoustique des sons dont l'analyse est donnée figure 6.3. Si ces trois sons donnent une sensation de hauteur suffisamment définie pour que nous puissions en chanter la note, seul le son de flûte, qui est périodique, permet l'appréciation de l'intervalle exact entre les deux sons : un « triton ». Pour les deux premiers sons, nous ne pouvons faire qu'une estimation, ce qui expliquerait que les auditeurs assimilent l'écart à l'intervalle voisin le plus familier : la quarte ou la quinte[4].

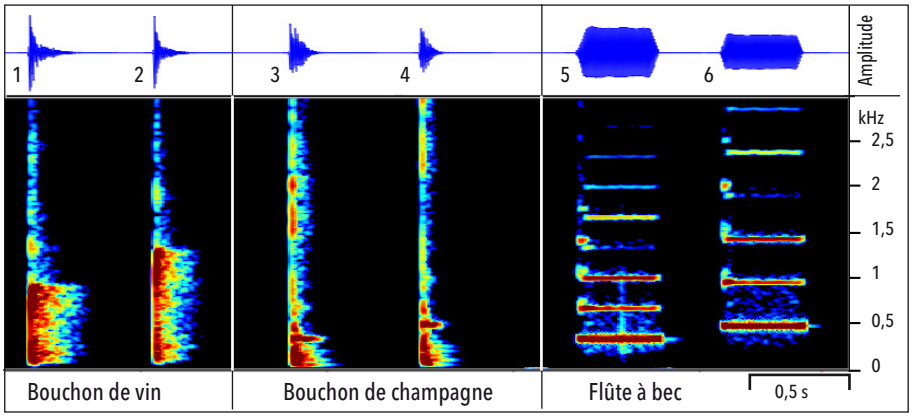


Figure 6.3 Analyse spectrale des trois sons formant un intervalle ascendant. Seul le troisième son, la flûte à bec, donne une sensation de hauteur précise car il est périodique. (Son 6.3).

La singularité du son périodique, déjà visible sur l'analyse sonagraphique de la figure 6.3, va se confirmer à l'issue de l'expérience suivante.

3. Les trois sons de cet exemple sont extraits de la collection « A » du chapitre 4, page 156.
4. Précisons pour les lecteurs peu familiers avec la musique que le triton – *diabolus in musica* – est un intervalle réputé difficile à entendre. Nous chantons plus aisément *do-fa* ou *do-sol* que *do-fa#*.

1.2.3. Expérience 3 : hauteur commune à différents sons

Cet exemple présente quatre sons nouveaux, tous différents (Son **6.4**).

Son 6.4 (6″)

Quatre sons de même « hauteur »

À la première écoute, les deux modalités perceptives se mêlent : reconnaissance de l'origine probable des sons (un bruit de débouchage – bouteille de champagne –, un sifflement, un souffle, une note de violon) et impression qualitative que ces quatre sons ont quelque chose de commun. À la deuxième écoute, l'impression se confirme : on peut siffler ou chanter une note commune à ces quatre sons (un *do*#$_4$).

Pourtant l'analyse spectrographique de la figure 6.4 est déroutante. Les trois premiers sons ont en commun une zone de fréquence dans laquelle l'énergie est renforcée (trait blanc vers 570 Hz), or cette zone fait totalement défaut (pointillés blancs) au 4e son, alors qu'il partage la même sensation de hauteur.

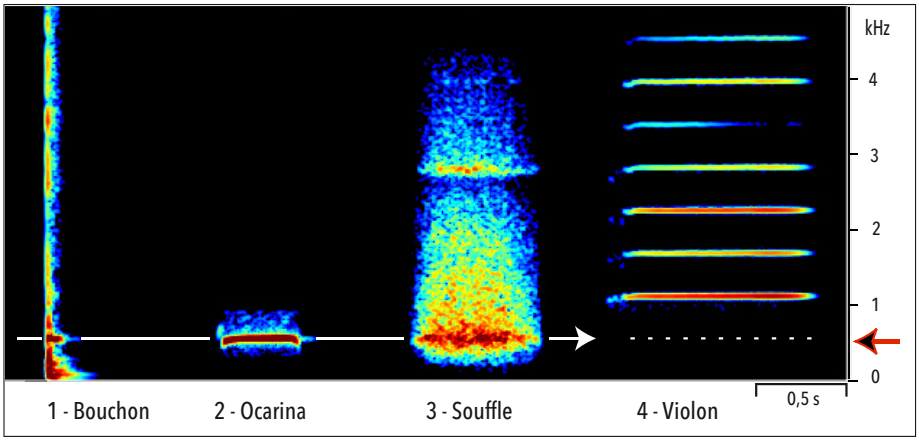


Figure 6.4 Analyse spectrale de quatre sons donnant la même sensation de hauteur. Hauteur spectrale (1 et 3) et hauteur tonale (2 et 4).

Détecter la périodicité

Comme dans l'exemple précédent, le quatrième son a une structure acoustique particulière : c'est un son entretenu périodique dont nous avons supprimé le premier harmonique. La hauteur n'est pas directement liée à la position du maximum spectral d'énergie, comme pour les sons 1 et 3, mais à notre capacité à détecter la périodicité qu'il y ait ou non de l'énergie dans la zone du spectre correspondant à la note entendue.

1.3. Bilan des trois expériences : la perception de hauteur et la structure acoustique des sons

Je ne puis admettre sans réserve la différence que l'on a coutume d'établir entre couleur sonore (timbre) et hauteur sonore. C'est en effet par sa couleur sonore – dont une dimension est la hauteur – que le son se signale.

Schoenberg, A., 1911, p. 516.

1/ **Hauteur spectrale**. À l'écoute de deux sons successifs de spectres différents, nous percevons une variation dont le sens ascendant ou descendant est lié au sens du déplacement global des maxima de l'enveloppe spectrale.

Hauteur spectrale et timbre. Pour des sons de spectre large (étendu), cette variation peut être qualifiée indifféremment de changement de hauteur ou de changement de couleur sonore, comme en témoigne le vocabulaire employé par les auditeurs pour décrire ce qu'ils perçoivent : son clair, aigu, brillant. Hauteur et timbre (dans le sens de couleur sonore) sont donc liées par notre capacité à coder et donc à comparer les maxima d'énergie du spectre (voir « codage tonotopique » chapitre 3, § 1.3.3, et chapitre 7, « hauteur spectrale », figure 7.3).

Hauteur spectrale et contour mélodique. Lorsque les sons ont un maximum spectral localisé dans une zone bien définie, la sensation d'une hauteur spectrale peut s'autonomiser et donner lieu à la perception de figures mélodiques plus ou moins bien définies. Analogue aux accents d'intonation de la langue, les contours mélodiques sont présents dans les musiques de percussion, notamment celles qui utilisent des tambours de hauteur variable (tambour d'aisselle, *ko-tsuzumi*, tom-tom, *Pakhavaj*, tabla basse).

La hauteur spectrale, de large bande ou localisée, s'applique aux sons de toute nature.

2/ **Hauteur tonale**. Le codage temporel, fondamentalement différent, ne s'applique qu'aux sons entretenus *périodiques***,** tels la flûte et le violon de nos expériences. La structure temporelle très particulière de ces sons donne lieu à une sensation de hauteur univoque et à une très grande finesse de discrimination entre les hauteurs.

Rappelons que seuls les sons entretenus, produits par la répétition du même mouvement vibratoire à intervalles de temps réguliers, sont périodiques. Ils ont la particularité d'avoir un spectre harmonique et constituent la plus grande part des sons produits par les instruments de l'orchestre (les instruments à vent et les cordes frottées). Nous verrons qu'on peut, dans certaines limites, assimiler à cette catégorie les sons *quasi périodiques* des cordes pincées et frappées, sans oublier que leur inharmonicité intrinsèque peut être la source de bizarreries perceptives qu'il nous faudra expliquer.

Dans la réalité de l'écoute musicale, les deux stratégies de hauteur spectrale et de hauteur tonale se combinent, entrent en concurrence et vont nous permettre d'interpréter la perception de la hauteur de la plus grande partie des sons de la musique, y compris ceux de la voix humaine, source de variations simultanées dans ces deux dimensions.

3/ **Hauteurs des sons inharmoniques**. La catégorie des sons *inharmoniques* est produite par la percussion de corps très peu amortis comme les cloches, les gongs, la timbale. Bien que ceux-ci aient été soigneusement accordés, leurs sons suscitent le plus souvent la sensation de plusieurs notes simultanées concurrentes, sensation qui varie d'un auditeur à l'autre. Pour ce type de sons, plusieurs stratégies d'écoute sont à invoquer : le codage temporel des fréquences des modes propres, les variations temporelles de telle ou telle composante, la zone fréquentielle dans laquelle se situent les composantes et surtout la recherche d'un modèle interprétatif que nous appelons « modèle harmonique ».

1.4. La dimension cognitive de la hauteur tonale : expérience de comparaison voix-sifflet

Les caractéristiques que nous venons d'énoncer ne doivent pas faire perdre de vue que la première réaction d'un auditeur à l'écoute d'un son est de lui attribuer une origine et un sens. La catégorisation de premier niveau se fait sur les sources sonores (voir chapitre 4, § 3.6 et § 4.).

Celles de la musique occidentale ont été sélectionnées et construites pour produire des sons de hauteurs différentes. Dans une catégorie instrumentale donnée (le violon, le piano), nous apprenons à différencier les sons selon leur hauteur. Certains instruments produisent des sons stables (piano, percussions), d'autres des sons variables (vents, cordes). Tous changent de contenu spectral, d'attaque et de dynamique lorsqu'on passe du grave à l'aigu de la tessiture. L'ensemble de ces variations, mémorisées, contribue à la reconnaissance de la hauteur des sons d'un instrument donné.

Parallèlement, la possibilité se développe de comparer les sons donnant la même hauteur sur des instruments différents, pour parvenir à l'abstraction notale évoquée plus haut. De même que pour la perception de l'intensité, l'élaboration de la catégorie « hauteur » se fonde donc sur une connaissance approfondie des sources et des changements de qualité associés aux variations de hauteur, connaissance qui s'élabore au cours de l'apprentissage musical, de même que nous apprenons à catégoriser les couleurs indépendamment des objets qui les portent.

En témoigne l'exemple sonore suivant, qui est déroutant, à moins que nous acceptions, comme cela est suggéré au début du chapitre, de considérer que la perception de hauteur n'est pas une donnée directe des sens mais le résultat de nombreuses opérations d'écoutes comparées pratiquées avec des sons accordés et réglés, c'est-à-dire ceux de nos instruments de musique.

L'expérience est très simple (Son **6.5**). Elle consiste à comparer deux sons donnant chacun une note : d'abord une voix, ensuite un son sifflé. La voix (V) est tout d'abord suivie du son sifflé dans l'aigu (S1), en deuxième lieu la même voix est suivie du son sifflé (S2) une octave plus bas que le premier.

Les questions posées sont :

Son 6.5 (8")

Voix et sifflet

Un des deux couples de sons (V, S1) ou (V, S2) produit-il une sensation de hauteur semblable ? En d'autres termes, y a-t-il un couple dans lequel les deux sons sont à l'unisson ? Si oui, lequel ? Sinon, pourquoi ?

Bien des auditeurs sont embarrassés pour répondre, et les avis sont très partagés. Les explications détaillées sont données dans l'encadré, page suivante.

Appariement de la voix et du sifflet

Cette expérience d'écoute, particulièrement déroutante pour les musiciens dont l'oreille est bien formée, peut s'expliquer par les remarques suivantes.

1/ Le son vocal et le son sifflé ont des contenus spectraux très différents. Le premier, le son vocal, possède plusieurs harmoniques alors que le second, le son sifflé, est pratiquement un son pur dépourvu d'harmoniques.

2/ Le sifflet, que nombre d'entre nous adoptent pour reproduire une mélodie n'entre pas dans la pratique musicale quotidienne pour l'évaluation de la hauteur sonore. Nous savons que le sifflet est « plus aigu » mais nous ne savons pas exactement à quel point.

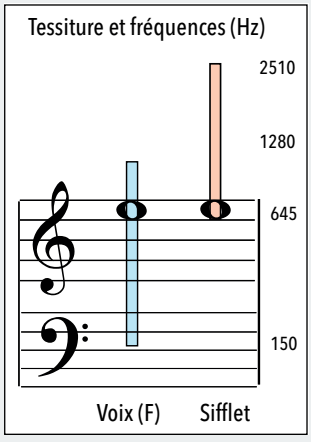


Figure 6.5 Les deux sources sonores : la voix et le sifflet, ont des tessitures différentes. Le mi_4 (645 Hz), sonne aigu pour une voix et grave pour un sifflet.

3/ La troisième explication, qui concerne la tessiture de chaque source, est selon nous la plus intéressante, car elle est généralisable à d'autres cas similaires. À l'écoute, le son vocal paraît aigu. Le mi_4 est effectivement un son de la partie supérieure d'une voix humaine. Au contraire, le son sifflé sur la même note, paraît grave pour un sifflet humain : il est difficile à produire et peu intense. Sur la figure 6.5 nous avons représenté par une barre verticale la zone de tessiture approximative des sons d'une voix féminine (en bleu) et celle d'un sifflet oral (en rouge). Un son de hauteur commune à ces deux sources se situe effectivement dans le grave de la tessiture du sifflet et dans l'aigu de celle de la voix.

Les deux couples du Son 6.5 correspondent aux deux premiers de la figure 6.6. Écoutez le Son **6.6** pour entendre le troisième couple (V, S3) à l'unisson véritable.[1]

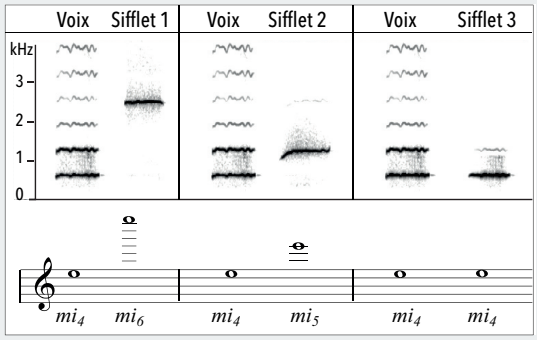


Figure 6.6 Analyse spectrale des sons (en haut) et notation musicale (en bas). L'intervalle entre les deux sons d'un couple voix-sifflet est successivement : deux octaves (V, S1) ; une octave (V, S2) ; l'unisson (V, S3 du Son 6.6).

Son 6.6 (15")

Comparaison de la voix et du sifflet avec unisson

1. Étrangement, le son S3, obtenu par transposition à l'octave inférieure de S2, est difficilement reconnu comme un son de sifflet, car il sort de la tessiture habituelle à cette production.

2. La hauteur des sons périodiques

2.1. Introduction : la singularité des sons purs

Il est habituel, lorsqu'on traite de la perception de la hauteur musicale, de commencer par l'étude des sons purs, les plus simples du point de vue physique. Il en va différemment pour un auditeur humain et cela pour deux raisons : d'une part, le son pur (ou sinusoïdal) se rencontre fort rarement dans la nature, c'est donc une catégorie sonore dont nous avons peu l'expérience[5] ; d'autre part, la sensation de hauteur d'un son pur est peu robuste. D'ailleurs, un musicien pourvu d'une excellente oreille absolue perd toute assurance lorsqu'on lui demande de prendre en dictée une mélodie de sons purs. En effet : la hauteur d'un son pur varie avec l'intensité : un crescendo fait baisser un son de 200 Hz et monter un son de 2000 Hz. La hauteur d'un son pur peut changer selon qu'on l'écoute (au casque) à l'oreille droite ou à l'oreille gauche, ce qui surprend vivement les musiciens. Enfin l'intervalle entre deux sons simultanés est beaucoup plus difficile à évaluer avec des sons purs qu'avec des sons riches en harmoniques (voir chapitre 8, § 2.2).

Au chapitre 3, nous avons rappelé les caractéristiques de l'oreille excitée par des sons purs : seuils absolus et différentiels, codages, limites perceptives. Nous allons maintenant découvrir que la perception de la hauteur musicale, qui permet d'identifier et de discriminer les hauteurs des sons complexes instrumentaux, se fonde souvent sur des repères autres que la seule fréquence des sons purs isolés entrant dans leur structure spectrale et fait appel à différents niveaux d'interprétation cognitive.

En bref, pour interpréter les réponses, il faut non seulement connaître les caractéristiques des oreilles des auditeurs, mais aussi et surtout la façon dont ils écoutent les sons. Certaines conduites d'écoute relevant de principes généraux (Gestalt, catégorisation) pourront être formalisées. D'autres, qui sont plus personnelles, correspondent à des habitudes d'écoute développées dans l'enfance dont nous avons rarement conscience. En cette matière, il n'y a pas de règle générale, pas de bonne réponse, mais des stratégies individuelles souvent fort complexes. C'est pourquoi la pratique des expériences d'écoute en groupe est particulièrement excitante lorsque les participants découvrent, en comparant leurs réponses, une disparité qui aiguise leur curiosité et mobilise le chercheur.

2.2. Du grave à l'aigu : les bornes de la perception de la hauteur tonale et la tessiture des instruments de musique

Limite grave. Le Son **6.7** fait entendre quelques notes jouées dans le grave du tuba. Cet exemple permet d'évaluer notre sensibilité aux mélodies très graves. Pour une partie des auditeurs, les sons entendus successivement se ressemblent, se répètent. D'autres perçoivent des changements difficiles à apprécier. Ces notes, qui ont des fréquences comprises entre 20 et 30 Hz, sont dans la zone limite de la fusion de succession des impulsions (voir chapitre 3, § 3.1), en pointillé sur l'axe de la figure 6.7. L'écoute des mêmes sons transposés à l'octave supérieure (Son **6.8**) permet de prendre conscience du point de basculement, lorsque la note devient repérable. Appliquée à la voix (Son **6.9**), une transposition à l'octave inférieure montre combien la limite de fusion est critique pour ce signal.

Son 6.7 (14″)

Tuba : 4 sons graves

Son 6.8 (53″)

Tuba : la séquence complète (12 sons) suivie d'une transposition *(voir page 281)*

Son 6.9 (22″)

Voix naturelle puis ralentie : fréquences divisées par 2 *(voir page 281)*

5. Pour produire mécaniquement un son quasi pur (dépourvu d'harmoniques), le plus simple est d'exciter un résonateur à la manière de la flûte, en soufflant très doucement sur le bord d'une bouteille. Parmi les oiseaux, certains comme le merle ont un chant comportant le plus souvent des sons purs : d'ailleurs on dit que le merle « siffle ».

Limite aiguë. À l'autre extrémité de l'axe des fréquences, au-delà de 4000 Hz, l'appréciation exacte de la hauteur d'un son devient très difficile voire impossible (voir chapitre 3, § 2.7.3). C'est la limite supérieure des notes du piano (environ do_7) et aussi le « plafond » des mixtures de l'orgue (voir figure 6.13A). Indépendamment des difficultés rencontrées dans la construction d'instruments de très petite taille, la raison principale de cette limite de tessiture est due aux limites du codage temporel de la hauteur.

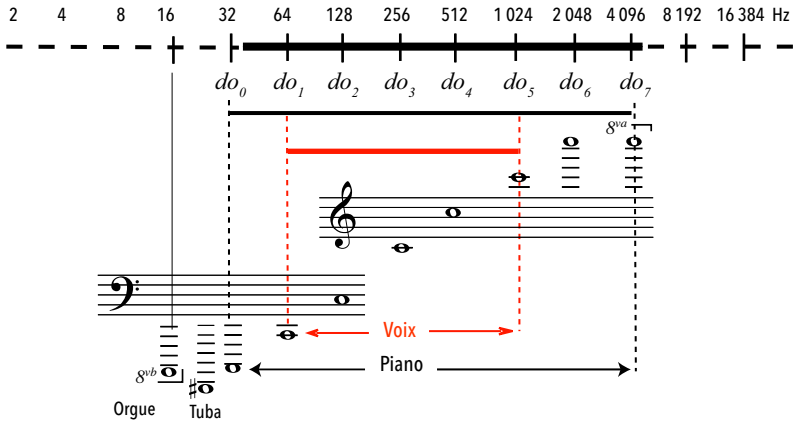


Figure 6.7 Les limites en fréquence du codage de la hauteur tonale (trait plein) correspondent à la tessiture de l'ensemble des instruments de l'orchestre. Notez la place centrale de la voix humaine.

Expériences de filtrage du basson (voir page 281)

Son 6.10a (18")

Filtre passe-haut

Son 6.10b (18")

Filtre passe-bande

Son 6.10c (18")

Filtre passe-bas

Son 6.10d (18")

Son original

Son 6.11 (5")

Basson, 1re note de chacune des séquences précédentes

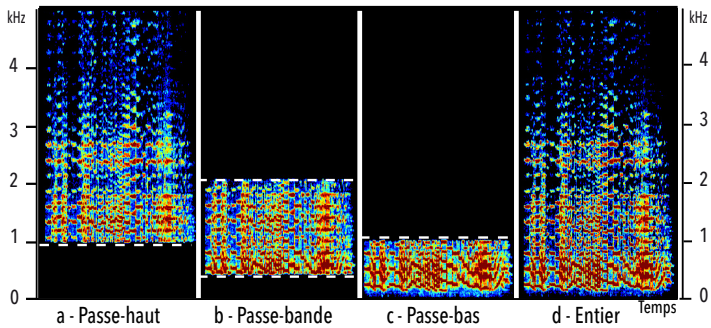
La figure 6.7 rassemble les résultats sous forme d'un graphique mettant en correspondance la suite des fréquences approchées[6] des *do* sur un axe gradué en échelle logarithmique, et la notation musicale des sons les plus graves et les plus aigus de quelques instruments. La tessiture tracée en rouge, centrée sur le do_3 (256 Hz), correspond à l'étendue moyenne des voix humaines réunies (hommes et femmes) et constitue la tessiture principale de la plupart des musiques traditionnelles.

2.3. La hauteur tonale et le filtrage du fondamental

L'opinion selon laquelle la perception de la hauteur est donnée par le premier harmonique, ou fondamental, est encore largement répandue. Dans les exemples suivants on peut entendre une phrase mélodique jouée au basson, répétée avec trois différents types de filtrages : passe-haut (Son **6.10a**) ; passe-bande (Son **6.10b**) ; passe-bas (Son **6.10c**) ; son naturel (Son **6.10d**).

Il est clair que la sonorité de l'instrument change mais que la mélodie n'est pas altérée. Le filtrage ne fait donc pas perdre la hauteur tonale, même lorsque les premiers harmoniques ont disparu (figure 6.8). Le fameux paradoxe du « fondamental absent » s'explique aisément lorsqu'on sait que le codage de la hauteur fondamentale d'un son périodique est temporel. Les formes d'onde de la partie supérieure de la figure 6.9, toutes différentes, montrent que la périodicité est conservée. Le Son **6.11** permet de saisir une seule note pour s'en assurer.

6. Le choix des fréquences approchées pour les *do* offre deux avantages : ce sont des nombres entiers et ils correspondent aux puissances de 2 que l'on spécifie lors des paramètres de l'analyse de Fourier. Au demeurant, ils correspondent aux *do* d'une échelle tempérée, plus basse d'environ un tiers de ton (la_3 = 430,5 Hz) que l'échelle de référence (la_3 = 440 Hz).



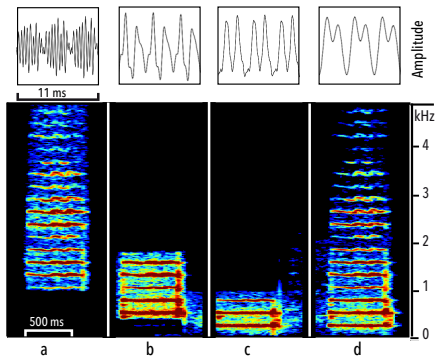


Figure 6.8 (à gauche) Filtrage d'une mélodie jouée au basson (Son 6.10). a) filtre passe-haut, b) filtre passe-bande, c) filtre passe-bas, d) son original

Figure 6.9 (à droite) Analyse de la première note (do_3 = 270 Hz) de chaque séquence filtrée. À la partie supérieure, le tracé du signal sonore pris dans un intervalle de temps de 11 ms correspond à 3 périodes du Son 6.11. La forme d'onde varie avec le filtrage mais la durée de la période est inchangée.

Cette situation d'écoute de sons graves filtrés « sans fondamental » est extrêmement banale. Nous la pratiquons quotidiennement au téléphone (voir chapitre 5, figure 5.3) ou lorsque nous écoutons la musique avec de petites enceintes. On sait que les premiers systèmes de reproduction sonore offraient une mauvaise restitution des basses fréquences, ce qui, joint à l'accentuation de la zone de fréquences 2000/3000 Hz, contribuait à donner un caractère nasal à la musique.

Trompes de chasse (voir page 281)

Son 6.12a (21″)

Son 6.12b (15″)

Son 6.12c (18″)

À titre d'exemple, on écoutera la reproduction de musiques de trompes de chasse (Son **6.12a** et Son **6.12b**) lues sur deux phonographes à aiguille ; le Son **6.12c** est un enregistrement professionnel restituant les basses fréquences.

2.4. Les liens entre le contenu spectral et la hauteur tonale perçue

Il n'est pas tout à fait exact de dire que le changement du contenu spectral n'a aucun effet sur la perception de la hauteur tonale. Voici trois exemples attestant de ce lien, au demeurant fort subtil, dont l'incidence est variable avec la tessiture et avec le type de sons et dont l'importance effective dépend beaucoup des habitudes d'écoute.

2.4.1. La richesse du spectre

Dans le Son **6.13** nous entendons la succession de trois sons isolés réalisés par synthèse. Ces trois sons ont très exactement la même période et pourtant de nombreux auditeurs ressentent un changement de hauteur, faible mais appréciable. Pour eux, le son paraît plus haut lorsqu'il est plus riche en harmoniques. C'est le même phénomène qu'il faut invoquer, notamment, lorsque deux groupes instrumentaux se répondent, l'un étant dans la salle et l'autre en coulisses. Ce dernier devra s'accorder à un diapason légèrement plus élevé pour paraître juste, afin de compenser la perte des harmoniques aigus due à l'absorption par les coulisses.

Orgue électronique (Cantor), sons de spectres différents (voir page 281)

Son 6.13 (11″)

2.4.2. La prédominance des harmoniques impairs

Les changements de spectre peuvent aussi affecter la sensation de justesse. Le jeu de la guitare, par exemple, exploite toute une gamme de variations de sonorité dues aux différents points de pincement de la corde (voir chapitre 7, Son **7.12**).

Pour certaines positions, des changements de hauteur peuvent se produire. C'est particulièrement le cas lorsque l'excitation est au milieu de la corde, ce qui produit une réjection de toutes les composantes de rang pair[7].

L'exemple **6.14** a été réalisé avec une longue corde montée sur un dispositif expérimental[8]. Il permet de comparer l'effet dû au point de pincement : en premier au milieu de la corde (Son **6.14a**), en second près de l'extrémité (Son **6.14b**). Les différences perçues entre les deux sons varient selon les auditeurs. Certains ressentent un changement de hauteur, de l'ordre d'un ton, d'autres, seulement un changement de timbre. Ce phénomène est pour l'instant sans explication.

Point de pincement d'une corde

Son 6.14a (7″)
Corde pincée au milieu

Son 6.14b (7″)
Corde pincée près du sillet

2.4.3. Le changement de diapason et son incidence sur la qualité sonore

Diapason et sonorité

Il s'agit cette fois de l'effet inverse : l'incidence du changement de fréquence sur les qualités globales du son. Une séquence musicale donnée, transposée d'un petit intervalle (inférieur au quart de ton), paraît changer de sonorité, comme si on en changeait le contenu harmonique (Son **6.15a**). Pour la grande majorité des auditeurs, le son semble plus terne lorsque le diapason baisse et plus brillant lorsqu'il monte. Le phénomène, qui est très sensible lorsqu'on passe rapidement d'un exemple à l'autre, reste à ce jour sans explication. L'enregistrement d'orgue de l'exemple **6.2.5** permet de comparer la même séquence jouée à l'orgue dans des conditions différentes de température, donc de diapason (voir les détails page 282). Cet effet dû au changement de diapason n'est sensible que lors d'un enchaînement à intervalle de temps très court et joue un rôle important lorsqu'on effectue des comparaisons fines de qualités entre deux instruments très semblables : il faut donc impérativement les accorder au même diapason.

Orgue et diapason

Son 6.15a (29″)
Voir page 282

Son 6.15b (25″)
Été

Son 6.15c (21″)
Hiver

2.5. La hauteur tonale et la hauteur spectrale : le contrebasson

Avant d'aborder ce paragraphe, il importe de se prêter à l'expérience suivante.

Son 6.16 (5″)

Contrebasson : cinq sons. Son à écouter muni d'un papier et d'un crayon.

Transcrivez la mélodie, soit sur une portée musicale, soit sous forme d'un dessin schématique.
L'exemple est court (5 notes de durée égale).

Pratiquée un grand nombre de fois avec divers groupes d'auditeurs, l'expérience fait apparaître deux types de réponses, assez également répartis. (Voir page 280.)

7. Le point de pincement, qui impose à la corde un grand déplacement, est un point de réjection de toutes les composantes spectrales qui y possèdent un nœud de vibration (voir chapitre 1, page 16).
8. L'inharmonicité n'est pas à invoquer ici puisqu'on compare deux sons produits par la même corde.

Il y a donc deux façons d'écouter cet exemple : soit en captant la hauteur fondamentale des sons, qui est très grave, soit en suivant le fil d'une mélodie formantique située dans le médium.

Dans cet étonnant exemple d'ambiguïté perceptive entre hauteur tonale et hauteur spectrale, deux mélodies coexistent et évoluent indépendamment l'une de l'autre. Pour qu'une telle ambiguïté soit possible, il faut, comme c'est le cas ici, que plusieurs conditions soient réunies (figures 6.10 et 6.59).

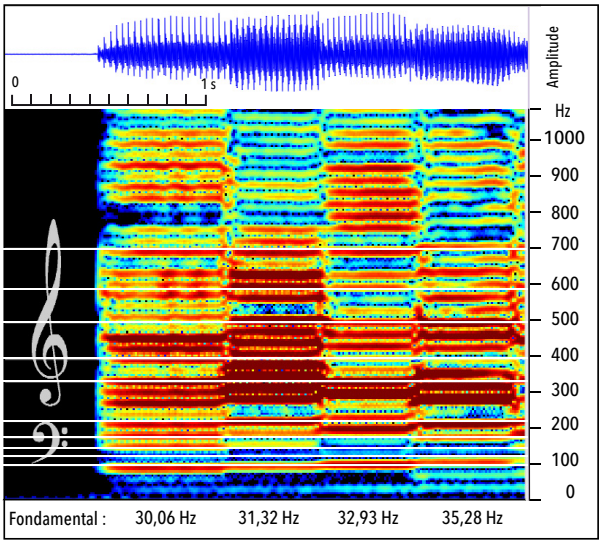


Figure 6.10 Hauteur tonale et hauteur spectrale des notes jouées au contrebasson. Analyse sonagraphique montrant les zones formantiques (en rouge foncé) et la faiblesse de l'énergie des premières composantes (bleu pâle).

- Les sons fondamentaux de l'instrument, situés dans une tessiture grave (entre 30 et 35 Hz), sont proches de la limite de fusion temporelle où la hauteur tonale est difficile à apprécier pour une partie des auditeurs.
- L'énergie des premiers harmoniques est extrêmement faible. Le premier harmonique très visible figure 6.10 est le troisième.
- Le spectre de l'instrument possède des formants situés dans la zone d'écoute dominante : entre 250 et 800 Hz (voir le chapitre 3, page 116). Le fait que ces formants varient d'une note à l'autre attire l'attention de l'auditeur : c'est le support de la mélodie spectrale (ou formantique).

Pour diversifier son écoute, l'auditeur s'efforcera d'entendre « l'autre mélodie », celle qu'il n'avait pas perçue spontanément, en remarquant que les deux modalités de l'écoute, spectrale et tonale, sont en compétition, l'une excluant l'autre, de façon analogue à la perception visuelle alternée d'un vase ou de profils de visages dans l'expérience de Rubin (chapitre 4, § 3.8). L'effet disparaît dès que le son progresse vers une tessiture plus haute ou lorsque les premières notes graves sont transposées à l'octave supérieure (Son **6.17**).

Son 6.17 (16")

Contrebasson, séquence complète Puis la même, transposée à l'octave supérieure

2.6. Les conflits entre hauteur tonale et hauteur spectrale : quelques sons paradoxaux

La dualité entre la hauteur spectrale et la hauteur tonale – dont nous verrons au § 2.7 qu'elle est intrinsèque à la voix humaine – a été exploitée systématiquement pour produire des sons aux effets surprenants.

Écoutons la gamme chromatique synthétisée en 1964 par Roger Shepard (Son **6.18**). Au bout d'une vingtaine de secondes pointe un doute, un étonnement qui grandit parce que les notes descendent interminablement...sans devenir plus graves !

Son 6.18 (36")

Son paradoxal de R. Shepard

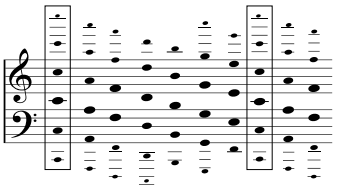


Figure 6.11 Représentation musicale de la descente d'une suite d'octaves dont les composantes ont une intensité constante dans la zone de fréquence de la portée musicale. (100-1000 Hz)

Les sons créés par Shepard pour cette expérience ont un spectre bien particulier qui n'est constitué que de composantes en rapport d'octave avec le fondamental (figures 6.11 et 6.13B). Par ailleurs l'énergie spectrale est concentrée dans une zone de fréquence qui reste fixe. En conséquence, la hauteur tonale descend, mais l'enveloppe spectrale est stable. De sorte qu'après avoir parcouru les douze sons de la gamme chromatique l'auditeur retrouve à l'octave inférieure, un son identique au son de départ, alors que dans un instrument réel, le spectre change avec la tessiture.

Pour trouver une issue à ce conflit perceptif chaque auditeur rétablit la sitation en remontant mentalement à l'octave supérieure : il « boucle » le trajet mélodique. Mais le remède n'est que provisoire, et au bout de douze notes il faut recommencer. Le malaise, ou l'excitation, sont analogues à ceux que l'on ressent en parcourant mentalement l'escalier de Penrose (figure 6.12).

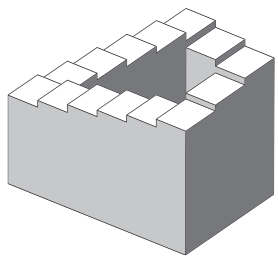
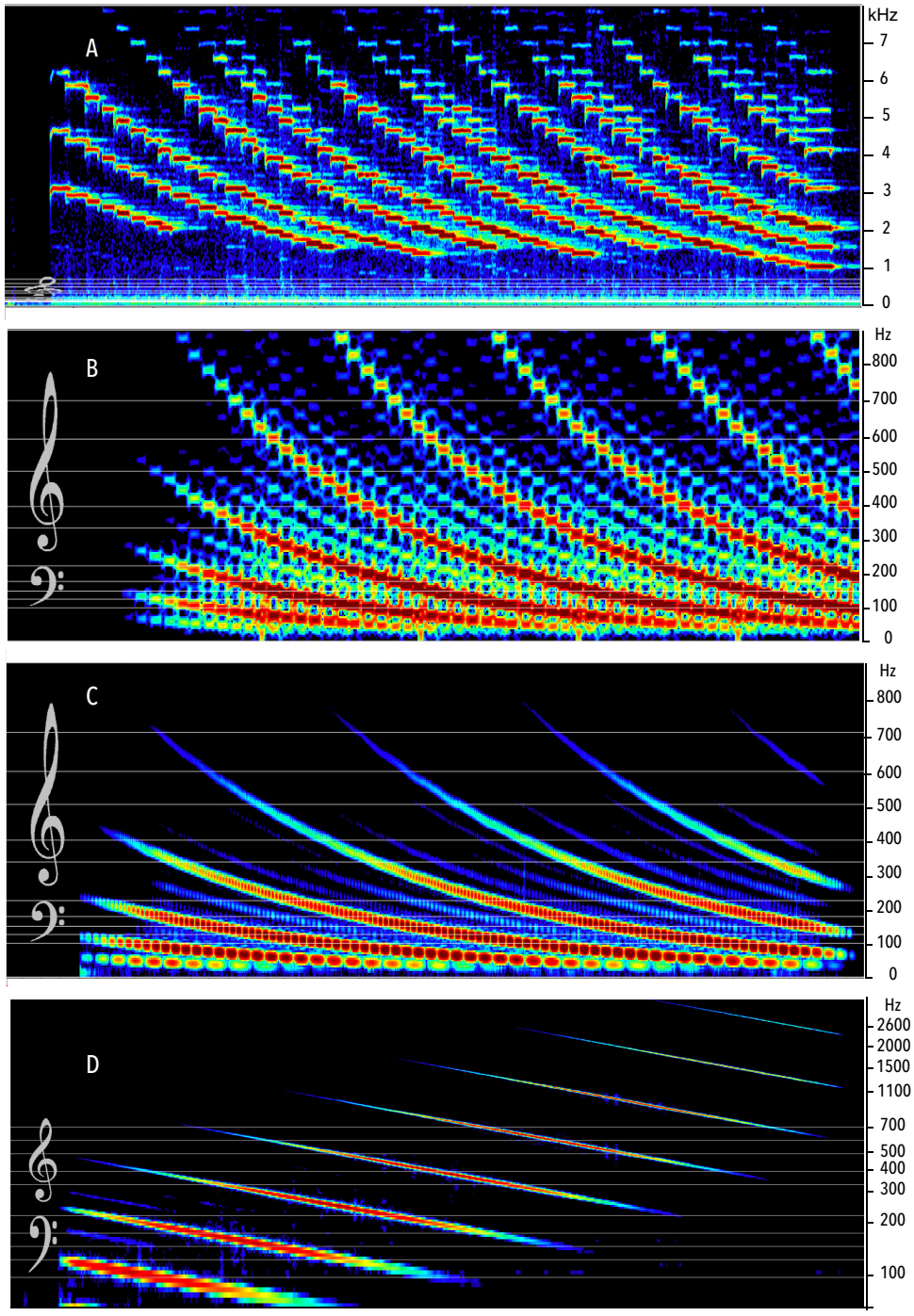


Figure 6.12 L'escalier sans fin de R. & L. Penrose, analogie visuelle du paradoxe perceptif produit par le son de R. Shepard.

D'après Penrose, L., Penrose, R., 1958, p.31-33.

L'effet d'une gamme descendante, jouée sur un clavier, qui donne l'impression de rester « sur place » est bien connu des facteurs d'orgue qui ont imaginé les « mixtures », dès la Renaissance (Son **6.19**). Constituées de rangs de tuyaux sonnant une suite de quintes et d'octaves au dessus de la note fondamentale à laquelle elles sont associées, les Mixtures du Plein-jeu (fourniture et cymbale) produisent un formant de « brillance » dans la zone 1000-4000 Hz. Comme dans l'exemple de Shepard, deux sons à intervalle d'octave, c'est-à-dire obtenus au terme d'une progression de onze sons sur le clavier, ont un spectre quasi identique et peuvent donc être entendus à la même hauteur. Voir l'analyse A de la figure 6.13.

Jean-Claude Risset a créé par synthèse différents sons paradoxaux combinant hauteur spectrale et hauteur tonale d'une grande qualité musicale. Le Son **6.20**, propose un glissando qui descend continuement ; dans le Son **6.21**, hauteur tonale et hauteur spectrale évoluent en sens contraire (figure 6.13, analyses C et D).



Son 6.19 (12″)

Orgue : gamme chromatique jouée sur la Cymbale seule (sans 8′)

Son 6.18 (36″)

Son de R. Shepard

Son 6.20 (42″)

Son dont la descente est infinie (Risset).

Son 6.21 (21″)

Son qui opère une descente continue en montant (Risset).

Figure 6.13 Quatre analyses de sons paradoxaux. A : gamme chromatique descendante jouée sur le registre de cymbale d'un orgue (Son 6.19). B : analyse du Son 6.18 de R. Shepard. C : analyse du Son 6.20 ; glissando mélodique sans fin, se poursuivant sans descendre. D : analyse du Son 6.21, descente mélodique infinie devenant de plus en plus « aiguë » par mouvements contraires entre la hauteur tonale (descendante) et la hauteur spectrale (ascendante). Pour ce son, l'échelle de fréquence est logarithmique (limite supérieure 4 500 Hz ; FFT : 8 192 pts).

2.7. La voix humaine : de la mélodie spectrale au chant harmonique[9]

2.7.1. La hauteur fondamentale et les formants vocaliques

La voix humaine est un exemple étonnant de « source » permettant de varier la hauteur tonale et la hauteur spectrale indépendamment l'une de l'autre. En effet, la production d'un son vocal s'élabore à deux niveaux. D'une part, les cordes vocales permettent de produire un son de hauteur fondamentale bien définie, c'est la mélodie de la voix, d'autre part, l'ajustement des cavités de résonance permet de doser le contenu harmonique du son émis, ce sont les voyelles.

Chacun de nous peut donc s'exercer à différencier la hauteur tonale et la hauteur spectrale sur sa propre voix.

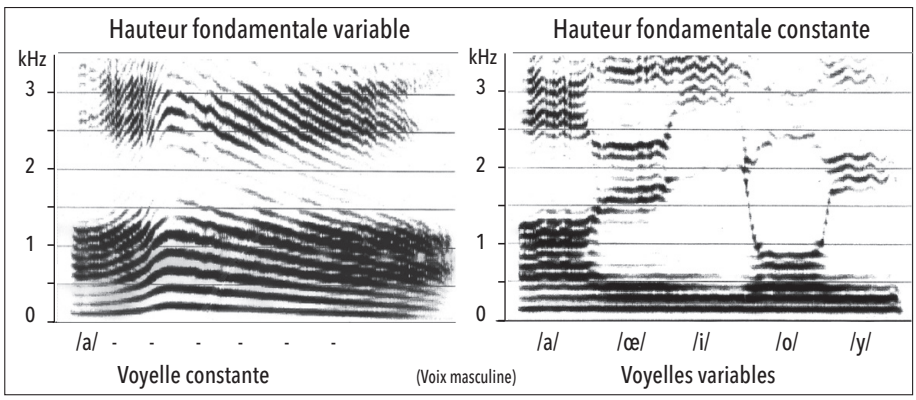


Figure 6.14 Deux séquences de voix masculine. À gauche : hauteur tonale variable (glissando ascendant - descendant) avec les formants de la voyelle « a », constants. À droite : hauteur tonale constante et variation des formants dans l'articulation des voyelles « a, e, i, o, u », notées sur la figure selon le code phonétique IPA.

Son 6.22 (10")

Exemple de la figure 6.14 (voix d'homme)

Voici un exemple : dire « a », et, sans changer la voyelle, faire glisser la voix en montant et en descendant. Ensuite, tenir la voix recto tono en articulant très lentement les voyelles « a, e, i, o, u ». La figure 6.14 montre l'analyse de ces deux exemples (Son **6.22)** et, notamment, les variations de la fréquence fondamentale sur le sonagramme de gauche, celle des formants spectraux sur le sonagramme de droite.

Certes, il n'est pas encore question de mélodie spectrale car, la voix étant l'« instrument » de la parole, notre écoute interprète prioritairement les variations de spectre comme des changements vocaliques. En d'autres termes, *il faut faire un effort d'attention pour « oublier » les voyelles,* c'est-à-dire passer de l'écoute sémantique (parole) à l'écoute de la mélodie spectrale (musique).

9. Que l'on désigne aussi par le terme chant diphonique.

Par exemple, si nous articulons plusieurs fois de suite et lentement les deux voyelles « ou, i » sur une note tenue (Son **6.23**), nous entendons assez clairement quelque chose qui monte et qui descend : c'est le mouvement spectral du deuxième formant (figure 6.15, gauche), déjà perceptible précédemment (figure 6.14).

Son 6.23 (7″)

Voyelles « ou-i-ou-i » avec une voix monocorde (voix féminine)

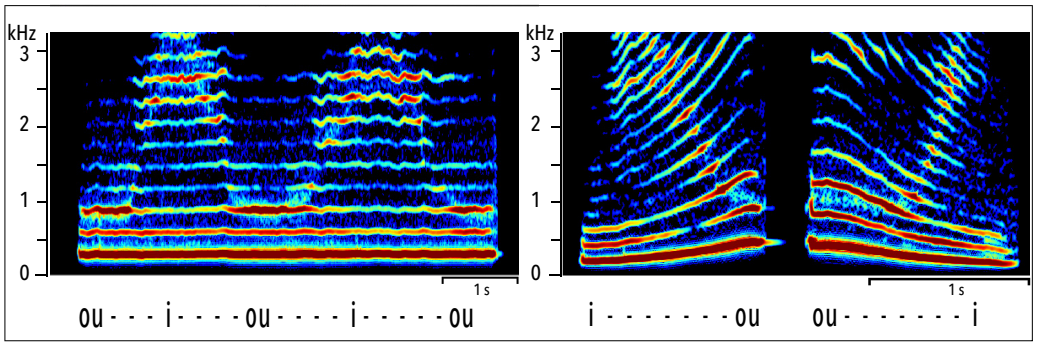


Figure 6.15 À gauche : sur une note tenue, dire lentement « ou-i-ou-i-ou », en écoutant le glissando spectral du deuxième formant vocalique. À droite : faire varier l'intonation de la voix en sens contraire du formant. Dire « i-ou » avec une intonation montante, puis « ou-i » avec une intonation descendante. De cette façon, la hauteur spectrale (le formant) et la hauteur tonale (l'intonation) varient en sens contraire l'une de l'autre (voix féminine).

On peut ensuite « jouer » à contrarier le mouvement du formant par une intonation variant dans le sens opposé (figure 6.15, droite, et Son **6.24**) et donc « croiser » une intonation ascendante avec un formant vocal descendant.

Son 6.24 (4″)

Variations de hauteur contraires entre l'intonation et le second formant (voix féminine)

Tous les exemples de musique traditionnelle du paragraphe suivant font entendre le mixage de mélodie tonale et de la mélodie spectrale qui utilise les résonances des différentes parties du système phonatoire humain, principalement le pharynx et la cavité buccale (pour une présentation plus détaillée de la voix, notamment le chant harmonique, voir le chapitre 9).

2.7.2. Les mélodies glottiques et les mélodies d'harmoniques

L'exemple du Son **6.25** marque une étape dans le changement perceptif. Cette fois nous n'entendons plus du tout de voyelle mais seulement une mélodie aiguë ascendante descendante dont les notes s'entendent distinctement lorsqu'on transpose l'exemple en le ralentissant (deuxième partie du Son 6.25). Cet exemple appelle plusieurs remarques.

Son 6.25 (20″)

Chant harmonique normal, puis ralenti et transposé à l'octave inférieure

- Lorsqu'un harmonique est considérablement renforcé par résonance, sa périodicité devient saillante perceptivement et il est perçu comme une hauteur indépendante. L'effet est connu depuis longtemps[10] mais, ici, le contraste entre l'immobilité de la hauteur fondamentale de la voix et les variations temporelles de la résonance fait totalement basculer l'écoute musicale en faveur d'une mélodie d'harmoniques doublée d'un bourdon grave monotone qui ne retient pas l'attention. Toutefois deux hauteurs sont bien en présence, d'où le terme de chant diphonique.

10. Le père Mersenne et Descartes ont signalé qu'il était possible d'entendre des sons plus aigus que la note fondamentale, à la quinte ou à la tierce, lors de l'écoute de sons graves d'orgue, et même de voix (voir chapitre 1, § 4.3).

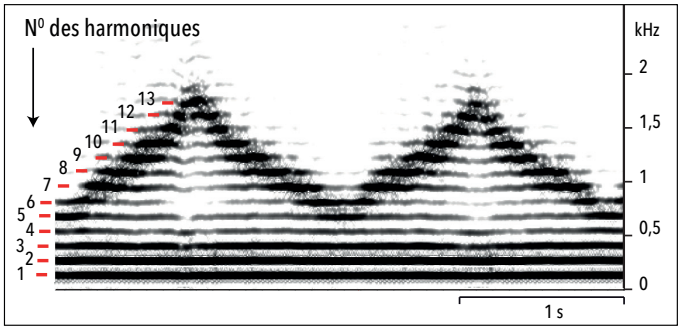


Figure 6.16 Mélodie spectrale produite par la technique du chant harmonique. La hauteur fondamentale (do_2 = 135 Hz) est fixe.

Son 6.25 (20")

Chant harmonique normal, puis ralenti et transposé à l'octave inférieure

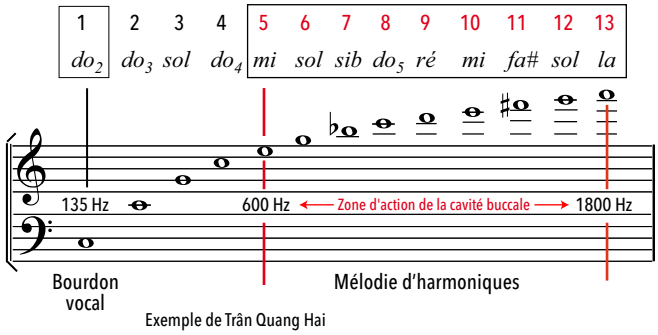


Figure 6.17 Harmoniques de la note fondamentale et zone d'action de la cavité buccale (2e formant vocalique) du Son 6.25.

- La technique vocale utilisée[11] consiste à ajuster la fréquence centrale du deuxième formant vocalique, dont la zone d'action varie entre 600 et 2000 Hz, tout en neutralisant le premier (figure 6.16). Les variations spectrales ne sont donc plus perçues comme des voyelles, mais bien comme une mélodie qui, dans cet exemple, évolue entre mi_4 et la_5 (figure 6.17).
- Une question se pose : la perception des mélodies harmoniques relève-t-elle de la hauteur tonale ou de la hauteur spectrale ?[12]. La réponse n'est pas simple, car cette perception varie selon la rapidité des changements spectraux[13], selon l'acuité de la résonance et surtout selon le type de source. En effet, parmi l'infinie variété des productions musicales existantes qui utilisent la bouche humaine pour sélectionner des composantes spectrales, certaines sont périodiques comme les cordes vocales, d'autres sont plus ou moins inharmoniques comme l'arc musical et la guimbarde. Ce qui caractérise toutefois ces mélodies, c'est le fait qu'elles nécessitent la sélection précise d'une seule composante spectrale, ce que l'on obtient au terme d'une pratique assidue. À cet égard, le Son **6.25**, témoigne de la maîtrise de son auteur, Trân Quang Hai !

11. Technique appelée « technique à deux cavités » par Trân Quang Hai.
12. C'est-à-dire d'un codage temporel ou d'un codage tonotopique ? (Voir chapitre 3, page 91)
13. Dans le Son 6.25, le glissando balaye dix notes à la seconde, ce qui est très rapide. Il faut ralentir le mouvement pour entendre des notes isolées.

2.7.3. Les mélodies d'harmoniques : quelques exemples musicaux

Les musiques de mélodies harmoniques sont connues depuis des temps immémoriaux. Leur développement a donné lieu à des techniques vocales d'une diversité et d'un raffinement impressionnants, associées parfois au jeu de la guimbarde (principalement en Asie centrale) ou à celui de l'arc musical (principalement en Afrique). Nous proposons ici l'analyse de quelques exemples particulièrement remarquables pour la perception de la hauteur. Des exemples similaires sont repris dans le chapitre 8 qui traite de l'analyse des échelles mélodiques. Pour une bonne compréhension des « formants » de la voix, allez au chapitre 9 où d'autres exemples sont également analysés.

1/ Arc musical (Afrique : pygmées ngbaka)

On peut pratiquer des mélodies spectrales buccales de deux façons différentes selon le formant vocalique exploité. L'effet perceptif change considérablement puisque le premier formant est utilisé entre 300 et 800 Hz et le deuxième formant entre 600 et 2000 Hz. Le fondamental de la voix doit donc doit être choisi en conséquence.

En effet, l'intervalle entre harmoniques consécutifs étant directement dépendant du rang des harmoniques sélectionnés – la tierce majeure apparaît entre les harmoniques 4 et 5 et le ton entre 8 et 9 – il faut que le bourdon (voix, guimbarde, arc) soit deux octaves plus bas que la limite inférieure du formant utilisé. Une solution alternative consiste à utiliser deux fondamentaux à intervalle d'un ton (voir chapitre 1, § 6). Dans l'exemple d'arc musical (Son **6.26**) de la figure 6.17, la mélodie principale due au deuxième formant évolue entre 700 et 1000 Hz, avec des incursions en glissando d'attaque jusque vers 1500 Hz (pour une analyse détaillée de cet exemple, voir chapitre 8, § 1.4.1).

Son 6.26 (51")

Arc musical à deux fondamentaux (Centrafrique)

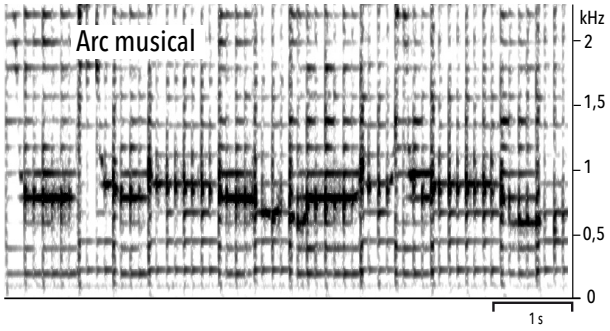


Figure 6.18 Mélodie jouée à l'arc musical en alternant deux sons fondamentaux de la corde à intervalle d'un ton.

2/ Voix et guimbarde (Asie centrale : tuva, style sygyt)

Dans l'exemple suivant (Son **6.27**), le musicien alterne la guimbarde seule (fondamental = 88,5 Hz soit un fa_1 + 30 cents) et la guimbarde doublée de la voix à l'octave supérieure, en utilisant le deuxième formant vocalique.

Son 6.27 (54")

Guimbarde et voix harmonique (Asie centrale)

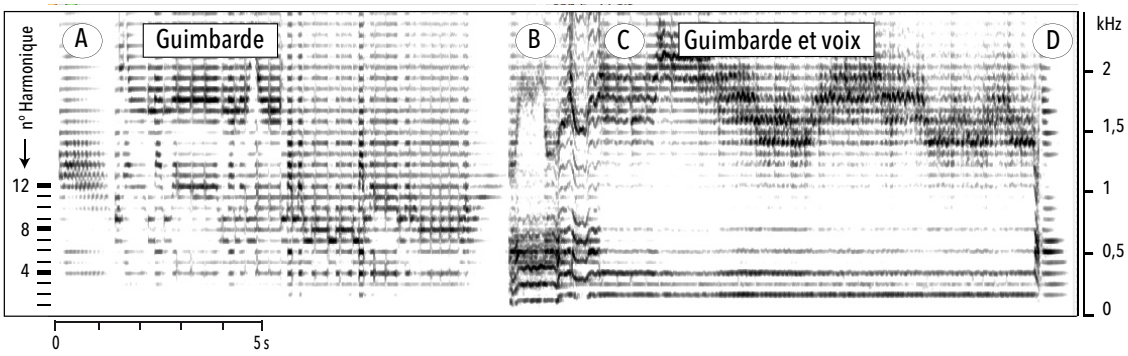


Figure 6.19 Guimbarde et chant harmonique (Son 6.27). A, guimbarde seule ; B, voix seule ; C, voix (style *sygyt*) et guimbarde simultanées.

Son 6.27 (54″)

Guimbarde et voix harmonique (Asie centrale)

Le jeu de la guimbarde seule (figure 6.19 A), est d'une grande complexité spectrale (Son **6.27**). Une ligne mélodique aux contours imprécis est perceptible entre 600 et 1 200 Hz.

À l'opposé, la mélodie harmonique de la voix s'impose sans ambiguïté dans la zone 1400-2100 (figure 6.19 C). Sur un fondamental fixe de 177 Hz, à l'octave supérieure du fondamental de la guimbarde (figure 6.20), la mélodie spectrale utilise les harmoniques 8 à 12 de la soure vocale (16 à 24 de la guimbarde). La rapide trémulation de la langue (8 Hz) produit une sorte de trille d'harmoniques.

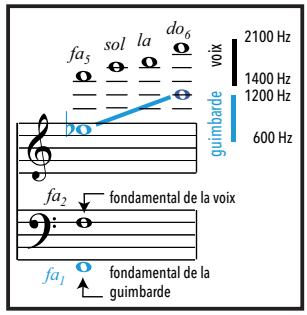


Figure 6.20 Notation musicale des tessitures mélodiques de la guimbarde et de la voix.

3/ Chant harmonique sur un fondamental très grave (Asie centrale : tuva, style kargiraa)

Son 6.28 (51″)

Chant harmonique kargiraa (voir page 283)

Voici un style de chant à l'écoute duquel on peut suivre simultanément l'écoute des voyelles et l'écoute d'une mélodie harmonique (Son **6.28**). En effet, à la différence de l'exemple précédent, les deux formants vocaliques sont présents et c'est au moyen d'ajustements très précis dans l'articulation des voyelles que le chanteur sélectionne l'harmonique désiré (figure 6.21). La zone d'action spectrale qui correspond au premier formant vocalique, se situe entre 400 et 800 Hz. Pour créer une mélodie entre les harmoniques 8 et 12, le chanteur pratique une technique vocale particulière qui produit un fondamental très grave, à 65 Hz (do_1). L'analyse détaillée de ce chant et de sa technique vocale est présentée au chapitre 9, § 3.3 et 3.4.

4/ Chant harmonique à deux notes fondamentales (Afrique du Sud : femme xhosa)

Son 6.29 (12″)

Chant harmonique à deux fondamentaux : femme xhosa (Afrique du Sud)

L'exemple suivant est remarquable à plus d'un titre (Son **6.29** et figure 6.22).

La technique vocale de la hauteur fondamentale dédoublée, similaire à celle du chanteur de l'exemple précédent, jointe à l'usage de deux notes fondamentales à un ton de distance (la_1 = 110 Hz ; si_1 = 122 Hz) permet à cette chanteuse de développer une mélodie pentatonique sur les harmoniques compris entre 3 et 6.

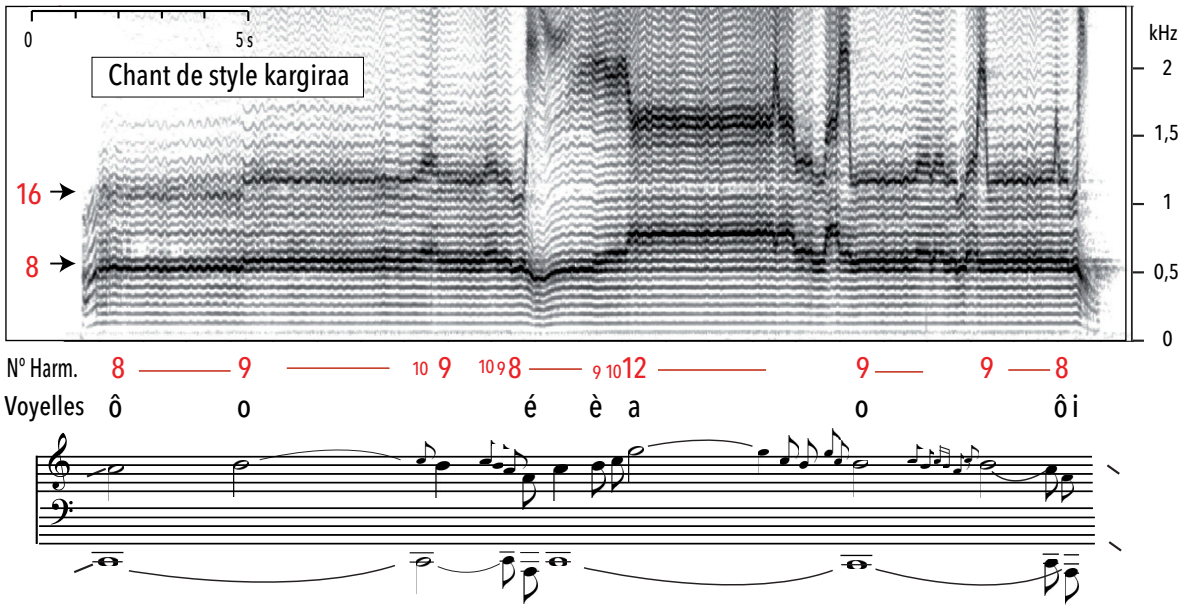


Figure 6.21 Chant diphonique de style kargiraa, dont la mélodie est produite par le premier formant vocalique. L'analyse détaillée de ce chant est reprise chapitre 9, § 3.3.

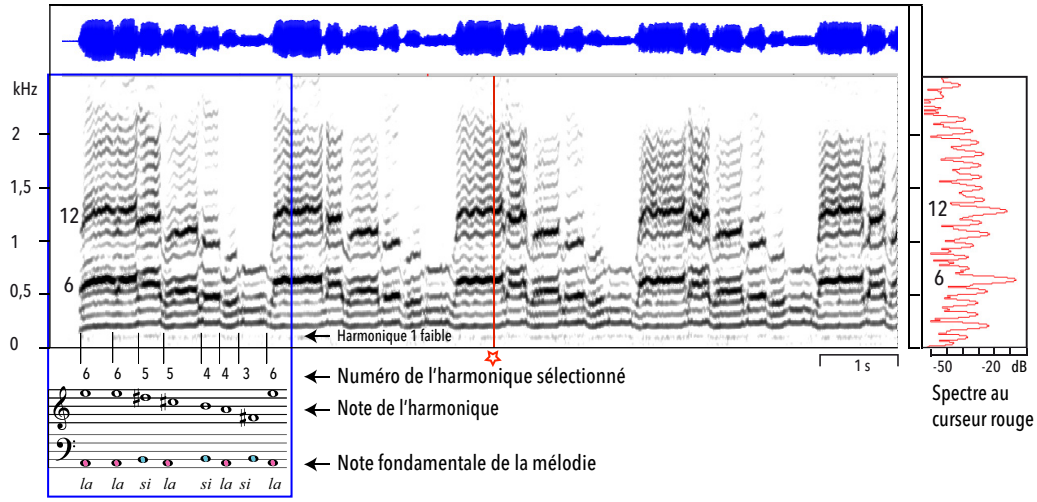


Figure 6.22 Chanteuse xhosa (Son 6.29). Mélodie spectrale descendante réalisée par le premier formant buccal dans la zone 350-650 Hz. Accord du deuxième formant à l'octave du premier. À droite de la figure, le spectre correspondant au curseur rouge met en évidence la saillance des harmoniques 6 et 12. Sous le sonagramme, notation musicale des deux fondamentaux la_1 et si_1 avec le numéro de l'harmonique renforcé par le premier formant.

La qualité particulière du timbre de la mélodie spectrale tient au fait que les deux formants vocaliques des voyelles utilisées, que l'on peut entendre comme « a, o, ô, ou », évoluent à l'octave l'un de l'autre. En effectuant un filtrage qui supprime le premier formant, on confirme le fait que c'est la mélodie située entre 350 et 650 Hz qui est perçue, renforcée par la doublure à l'octave.

Unisson de la hauteur harmonique et de la hauteur fondamentale (Tibet : chant rituel)

Cet exemple exceptionnel nous a été signalé par l'éthnomusicologue Mirelle Helffer. Au début du chant, on peut entendre un soliste dont la voix très grave[14] (environ 66 Hz, do_1), est émise en technique dédoublée. Il chante la voyelle « o » avec un harmonique 5 (mi_3) extrêmement intense, puis il est rejoint par le chœur. Vers la deuxième minute, le chœur entreprend seul une montée progressive et lente qui débute sur fa_1 (89 Hz) et dure six minutes. Peu après une nouvelle entrée du soliste (chrono) le chœur des moines atteint 325 Hz (mi_3), réalisant ainsi une fusion étonnante entre la note fondamentale du chœur et celle de l'harmonique 5 du soliste, fusion visible sur la figure 6.24.

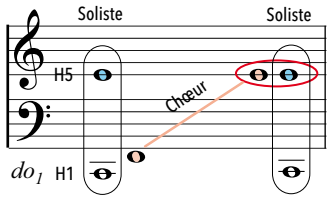


Figure 6.23 Notation schématique du son 6.30. Les notes bleues correspondant à l'harmonique 5 du soliste.

Son 6.30 (3'40)

Chant rituel tibétain : unisson de la hauteur fondamentale et de la hauteur harmonique (montage expérimental de démonstration)

Le montage raccourci de la séquence du Son **6.30** permet d'entendre brièvement ce phénomène exceptionnel mais ne dispense pas d'écouter l'exemple réel dans lequel la lente montée participe d'une tension croissante qui trouve son accomplissement dans l'unisson final (la durée de ce fragment est de 9' 28''). Voir la notation musicale schématique sur la figure 6.23.

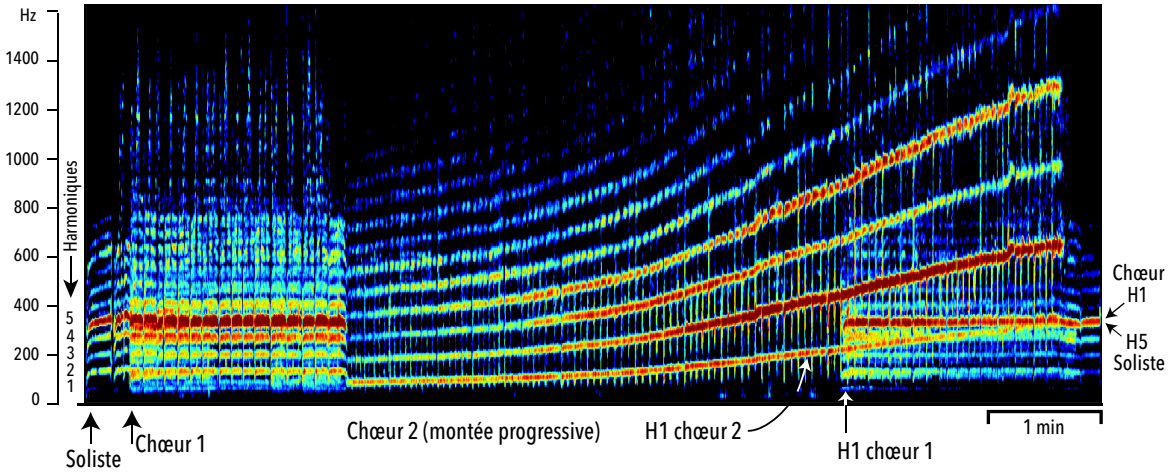


Figure 6.24 Chant tibétain. Unisson de hauteur fondamentale (chœur) et de hauteur spectrale (harmonique 5 du soliste). Son 6.30.

Les différents exemples présentés ici, illustrent les immenses possibilités de la voix humaine, avec laquelle on peut « jouer » des mélodies tonales, des mélodies spectrales, des mélodies vocaliques ou quasi sifflées.

14. La fréquence fondamentale fluctue légèrement entre 66 et 69 Hz, donc entre un do_1 haut et un $réb_1$ bas. Nous prenons le parti d'adopter do_1 comme base pour l'analyse et la notation musicale.

2.7.4. En résumé

Les musiques harmoniques, dont la mélodie est due à la sélection de composantes spectrales donnent lieu à des productions musicales d'une grande variété. Elles sollicitent l'écoute de différentes façons, selon que l'accent est mis sur les voyelles ou sur les fréquences des formants et les mélodies spectrales. On peut y rencontrer :

- des productions musicales purement mélodiques (sans voyelles) : arc musical, guimbarde, chant sygyt ;
- des exemples associant mélodie formantique et timbre vocalique : chant kargiraa ;
- des exemples de mélodies spectrales produites par deux formants en rapport d'octave : chant xhosa ;
- un exemple exceptionnel de fusion entre la hauteur spectrale due au premier formant et la hauteur tonale : chant rituel tibétain.

3. La hauteur des sons apériodiques

3.1. Introduction

Dès que le son n'est plus entretenu, la perception de hauteur devient rapidement très complexe à étudier. En effet, les composantes spectrales correspondant aux modes propres ne sont plus synchronisées sur celle d'un mode donné et, a priori, chacune d'entre elles peut être perçue isolément. Plusieurs hauteurs ou « notes » sont donc en concurrence. Les capacités sensorielles des auditeurs et leurs habitudes d'écoute étant différentes, on peut s'attendre à une grande variabilité dans la perception de la hauteur.

Par ailleurs, les sons apériodiques – à l'exception de ceux réalisés par synthèse – sont produits par une seule impulsion. Leur durée correspond à l'extinction de la vibration. La dimension temporelle prend alors une place importante dans l'appréciation perceptive puisque le son évolue constamment pendant toute son existence. Certaines composantes s'éteignent plus vite que d'autres, certaines peuvent même apparaître.

Remarquons que pour une bonne partie des sons apériodiques, ce n'est pas la seule perception d'une hauteur bien définie qui est recherchée, mais une heureuse conjonction entre les caractéristiques temporelles et les caractéristiques spectrales, ce qui constitue le timbre même du son. On sait que le coefficient d'amortissement (voir Glossaire) – qui conditionne la durée d'extinction du son – dépend en grande partie de la nature des matériaux mis en vibration et que les fréquences des modes propres, donc celles des composantes spectrales du son, sont conditionnées à la fois par la géométrie de la structure et par les caractéristiques des matériaux.

Entre le son d'un *wood-block*, de durée très brève et de hauteur spectrale peu définie, et celui d'un gong dont l'évolution temporelle produit un foisonnement sonore complexe en évolution constante pendant au moins 30 s, il y a place pour une très grande diversité de sons. Dans ce chapitre, nous ne retiendrons que les sons « accordés », ceux dont la notation musicale peut être représentée par une note sur portée.

Plus que jamais, il importe d'écouter et de faire en sorte que les auditeurs puissent comparer leurs avis. Lorsqu'une interprétation est proposée seule la « preuve » par la synthèse, qui permet de modifier les sons, est recevable. Nous en ferons largement usage. Voici, pour commencer, une expérience intrigante.

Le son du piano est inharmonique

3.2. Un cas particulier : la quasi-périodicité des sons de piano

Imaginons un instant que notre oreille a perdu la sensibilité aux sons graves et ne capte plus que les fréquences supérieures à 1 200 Hz.

Voici l'effet que nous ferait l'écoute de deux sons de piano de même fréquence fondamentale dont il ne reste plus que la partie aiguë du spectre (Son **6.31**) : nous percevons un intervalle ascendant entre les deux notes. En restituant la totalité du spectre (Son **6.32**) nous sommes surpris d'entendre la même note ! Le deuxième son paraît peut-être plus « brillant » mais pas vraiment plus haut du point de vue de la justesse[15].

Effet de l'inharmonicité sur la hauteur des sons de piano (voir page 283)

Son 6.31 (5")

Sons filtrés A et B

Son 6.32 (5")

Sons originaux A et B

Cette expérience réalisée par Taro Mori (2000) est constituée d'un son réel de piano échantillonné (B sur la figure 6.25) à partir duquel un nouveau son de piano totalement harmonique a été obtenu (le son A).

Nous pouvons donc comparer, sur la même fréquence fondamentale[16], deux sons ayant l'attaque et l'enveloppe dynamique d'un son de piano, mais une structure spectrale différente. Par comparaison aux harmoniques du son A, les partiels du son B sont légèrement plus aigus et cet écart croît avec le rang de la composante.

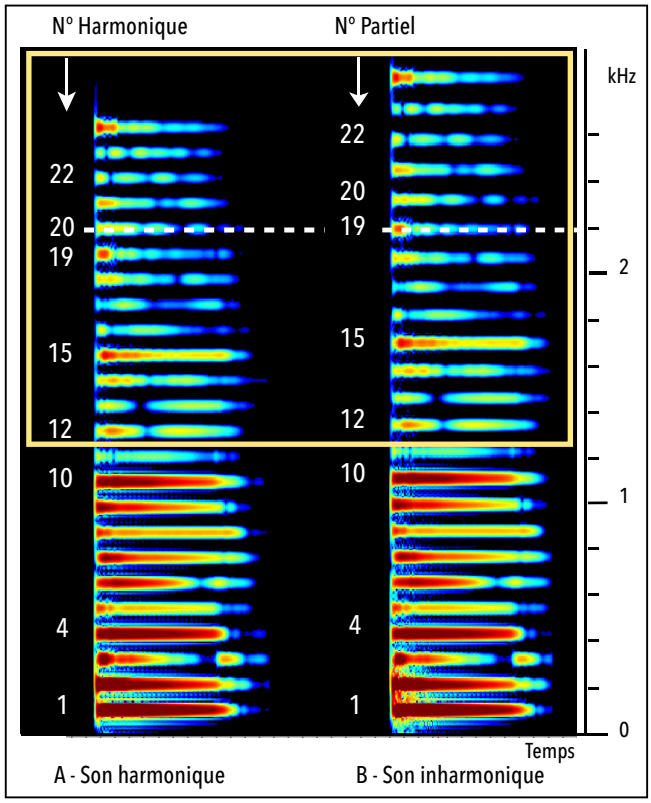


Figure 6.25 Piano. Comparaison des spectres d'un son périodique (à gauche) et d'un son réel de piano, inharmonique (à droite).

15. Cette expérience est à rapprocher de celle du § 2.3 où une mélodie jouée au basson (dont le son est périodique) n'est pas modifiée par le filtrage.
16. L'harmonique 1 du son A et le premier mode vibratoire du son B ont la même fréquence.

La figure 6.25 montre que le partiel 19 du son réel de piano (B) est à la même fréquence que l'harmonique 20 du son périodique (A). L'intervalle entre ces deux sons (20/19), est proche d'un demi-ton (voir annexe A). Il correspond à l'intervalle entendu dans l'exemple **6.31**, lorsque seuls les partiels 12 à 24 sont présents.

Il peut sembler étonnant, voire choquant, de découvrir que les composantes du son du piano, instrument si familier et si répandu dans la littérature musicale, ne soient pas des harmoniques ! Le piano est l'instrument de l'éducation musicale de bon nombre d'enfants, et celui sur lequel les professeurs font habituellement la démonstration des « harmoniques » vibrant par sympathie. En fait, notre oreille est tolérante.

Comme nous l'avons affirmé à plusieurs reprises dans le premier chapitre, seuls les sons entretenus ont des composantes harmoniques. Le son d'une corde de piano, frappée, est inharmonique. Mais tant que l'inharmonicité reste faible (ß <0,0002)[17], le signal temporel présente une pseudo-périodicité qui permet de l'assimiler perceptivement à un son périodique (figure 6.26). Cette inharmonicité est d'ailleurs ressentie et interprétée comme une caractéristique fondamentale du timbre du piano. Le problème de l'appréciation exacte de la hauteur d'un tel son, apériodique, se pose lorsqu'il s'agit d'accorder très précisément l'instrument (voir le chapitre 8, § 3.3.2).

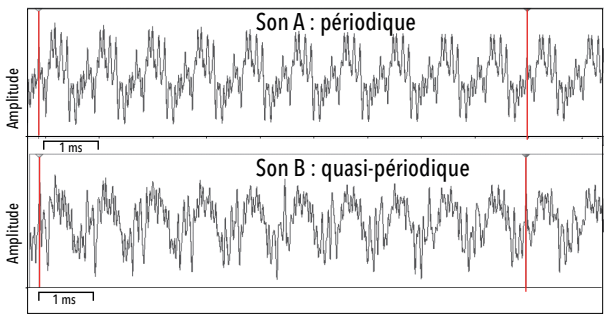


Figure 6.26 Comparez attentivement la forme de chacune des 10 oscillations des deux sons. La forme de l'onde A est très stable d'une période à l'autre. Celle de l'onde B change d'enveloppe de l'une à l'autre, tout en montrant une certaine stabilité de l'ensemble. Les deux repères rouges (à l'intérieur desquels on compte 10 cycles) se correspondent assez bien.

Dans le registre grave, l'inharmonicité est beaucoup plus perceptible. L'oreille étant plus sensible aux composantes aiguës, elle prend en compte les partiels élevés pour interpréter la hauteur, or ce sont les plus inharmoniques. Là aussi l'habitude (et en particulier la référence à son instrument personnel) joue un rôle important. À titre de démonstration, on pourra faire l'expérience suivante : jouer simultanément do_0 à la main gauche et $do\#_6$ à la main droite, puis $do\#_0$ et $ré_7$. Ce qui, à la vue des touches jouées, paraît discordant, ne choque pas l'oreille d'un auditeur non prévenu. En effet, les composantes de rang élevé (16 à 32) des sons graves étant, comme nous venons de le voir, plus hautes d'environ 1/2 ton, se retrouvent à l'unisson des fondamentaux des sons aigus. Notons qu'il vaut mieux faire l'expérience à deux, une personne joue et l'autre écoute, car le fait d'apparier des sons a priori dissonants sur le clavier induit une écoute biaisée chez la personne qui joue.

17. Le coefficient d'inharmonicité ß dépend de la raideur des cordes. Il est plus grand sur les pianos droits (cordes courtes et plus raides) que sur les pianos à queue.

En conclusion, le groupe des instruments à cordes frappées (piano) ou pincées (harpe, guitare, luth) constitue une catégorie à part pour la perception de hauteur et pour tout ce qui concerne la justesse et l'accordage. L'inharmonicité, dépend prioritairement de la qualité des cordes. Grâce au soin apporté à leur fabrication et aux progrès des techniques de filage qui ont permis d'alourdir les cordes graves tout en leur conservant une souplesse suffisante, l'inharmonicité reste est faible dans les instruments de l'orchestre classique.

3.3. La hauteur due à une composante dominante : le diapason à fourche et le marimba

Le diapason à fourche

Le diapason à fourche est un paradoxe : c'est un système vibrant non entretenu, donc apériodique[18]. Or, il est traditionnellement l'instrument de référence pour l'accordage des instruments de musique, en raison de ses qualités pratiques (stabilité en fréquence, portabilité), mais aussi pour ses qualités spectrales.

Son 6.33 (4")

Diapason à fourche

La figure 6.27, déjà proposée chapitre 1, page 16, représente l'analyse du son d'un diapason à fourche que l'on frappe et que l'on pose rapidement sur un support de bois. (Son **6.33**). Le son global est effectivement apériodique, c'est-à-dire que les composantes des divers modes ne sont pas liées entre elles par une période d'entretien.

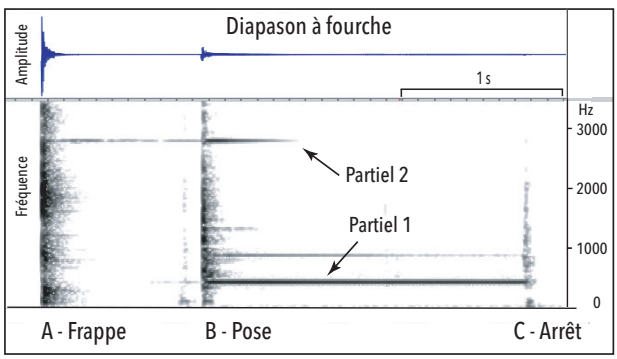


Figure 6.27 Analyse sonagraphique du son d'un diapason à fourche. Le diapason est d'abord percuté (A) puis posé sur un support de résonance (B). On voit les fréquences des deux premiers partiels : 440 Hz et 2 800 Hz.

À l'usage, le partiel 1 rayonne efficacement lorsqu'on couple le diapason à une boîte de résonance ou lorsqu'on l'approche de l'oreille. Le partiel 2 qui est très aigu (environ 2 800 Hz), s'amortit trop rapidement pour perturber l'écoute de la note recherchée qui est le la_3 : il n'est pas pris en compte dans la perception de la hauteur du diapason.

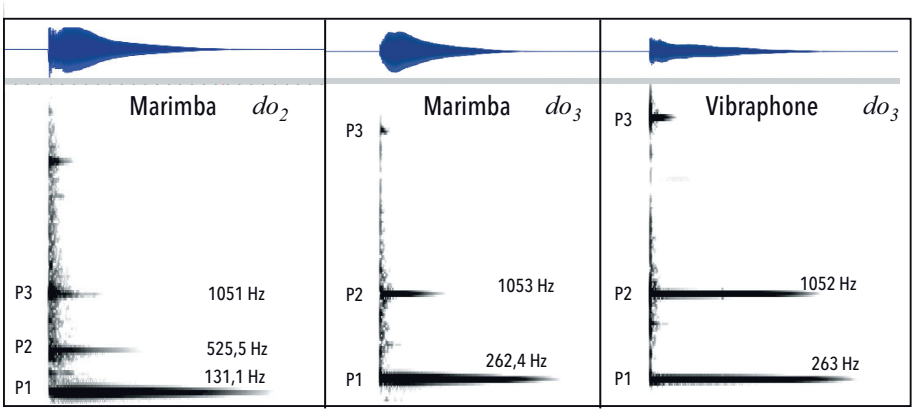
C'est donc la seule fréquence du partiel 1 qui donne le « la_3 ».

Marimba et vibraphone

Bien que fort différent dans sa structure et par son matériau, le marimba peut être rapproché du diapason en ce qui concerne la perception de la hauteur. Écoutons le son d'une lame de marimba (Son **6.34**). Dans cet instrument aussi, c'est le partiel du premier mode vibratoire qui donne la note (attention : si vous écoutez cet exemple sur les enceintes de votre ordinateur, il se peut que vous n'entendiez pas le vrai fondamental de la lame, à 131 Hz, mais seulement le partiel 2 qui est deux octaves

18. Cet instrument est apparu au XVIIIe siècle, mais son invention, attribuée à John Shore, n'est pas connue avec certitude.

au-dessus). L'instrument est conçu de telle sorte que l'intensité et la durée du premier mode de la lame sont renforcés par l'action d'un tube, situé au-dessous, tube dont la fréquence de résonance est accordée sur celle du premier mode de la lame.



Son 6.34 (3")

do_2 joué au marimba (à écouter dans de bonnes conditions)

Figure 6.28 Analyses comparées de deux sons joués au marimba (do_2, do_3) et du même do_3 joué au vibraphone. Remarquez la forme caractéristique de l'enveloppe d'amplitude du marimba (en bleu). Le gonflement qui suit la frappe est produit par la mise en route du tube accordé sur la fréquence fondamentale du premier partiel P1. (Mesures des partiels filtrés avec Praat)

Dans l'exemple sonore **6.35**, on peut comparer le son d'une lame de marimba en bois et celui d'une lame de vibraphone en métal. L'amortissement du métal étant beaucoup plus faible, le vibraphone a une sonorité plus claire et une tenue plus longue. Le son du marimba paraît moins percussif car, d'une part, la frappe est produite avec une mailloche douce et, d'autre part, la résonance du tube situé sous la lame se met en route avec un décalage temporel qui produit un effet de crescendo à l'attaque, très visible sur la forme de l'enveloppe dynamique à la partie supérieure de la figure 6.28.

Son 6.35 (6")

do_3 joué au marimba, puis au vibraphone

L'accordage des modes supérieurs des instruments à lame varie selon les facteurs. On voit sur notre analyse que les partiels 2 et 3 de ces instruments sont accordés respectivement sur des fréquences voisines des octaves double et triple du partiel 1, ce qui contribue à une meilleure définition de la hauteur tout en participant de la qualité sonore de l'instrument. L'emploi de baguettes plus ou moins dures permet aussi d'agir sur le filtrage des modes supérieurs, plutôt indésirables dans le grave de la tessiture.

3.4. Le modèle harmonique : cloches et timbale

La grande majorité des sons produits par choc n'est ni quasi périodique, comme le son de piano, ni pourvue d'une composante dominante qui donne la hauteur, comme le son de marimba. Leur spectre comporte plusieurs composantes inharmoniques, quelquefois sans rapport direct avec la hauteur perçue. Lorsque le son donne une hauteur bien définie à l'oreille, c'est le plus souvent sur la base du modèle harmonique.

3.4.1. Un exemple paradoxal : la cloche tube

La cloche tube

Écoutons le Son **6.36**. Il s'agit d'un tube métallique employé à l'orchestre dans le jeu de cloches (aussi appelé carillon tubulaire, *chimes* en anglais). La note de ce tube peut être chantée ou trouvée sur un piano : c'est un do_3. Pourtant, elle n'a pas d'existence physique et aucune des composantes n'a pour fréquence un multiple entier du do_3.

Son 6.36 (3″)

Note do_3

L'analyse spectrale de ce son est donnée figure 6.29 ; elle est suivie de celle du son pur qui correspond à la hauteur perçue : do_3 = 263,2 Hz (Son **6.37**).

Son 6.37 (6″)

Note do_3 puis son pur à la fréquence du do_3

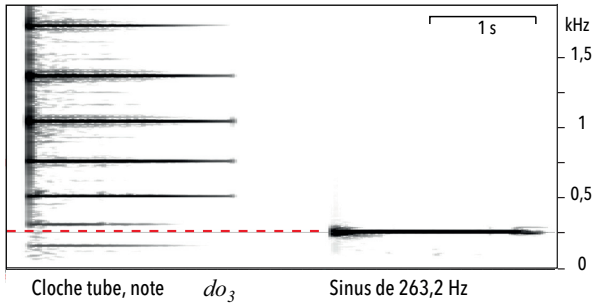


Figure 6.29 Sonagramme du Son 6.37. Successivement le son du tube et celui d'un son pur à la fréquence du do_3. Aucun des partiels du tube ne correspond à la note entendue. À droite, mesures comparées des harmoniques d'un do_3 et des six partiels de la cloche tube qui joue la note do_3.

Tableau 1. Les partiels d'une cloche tube

Nos	1	2	3	4	5	6
Partiels de la cloche tube (Hz)	318 P2	**520 P3**	**772 P4**	**1 064 P5**	1 394 P6	1 757 P7
Harmoniques d'un do_3 (Hz)	264 do_3	**528 do_4**	**792 sol_4**	**1 056 do_5**	1 320 mi_5	1 584 sol_5

Son 6.38 (35″)

Analyse auditive des partiels du tube (voir page 283)

Le tableau 1 montre les fréquences des partiels du tube et celles des harmoniques du do_3. En examinant le tableau et en écoutant les composantes isolées du tube (Son **6.38**), on remarque trois notes (P3, P4, P5), voisines respectivement de do_4, sol_4 *et* do_5. Ces trois composantes suffisent à suggérer un pseudo-spectre harmonique. Si on filtre le son de cloche tube de façon à supprimer les partiels supérieurs (au-dessus du partiel 5 à 1 064 Hz), on vérifie que la hauteur perçue ne change pas (Son **6.39**).

Son 6.39 (7″)

Partiels supérieurs filtrés puis son pur (voir page 284)

Cette expérience est un exemple de l'attribution d'une hauteur sur la base du modèle harmonique. Dès que l'auditeur peut approximer quelques composantes à une suite d'harmoniques, il « entend » une fondamentale.

Dans le cas présent, les composantes s'écartent notablement d'une série harmonique. L'intervalle P4/P3 est plus petit qu'une quinte (1,48 au lieu de 1,5). L'intervalle P5/P4 est plus grand qu'une quarte (1,37 au lieu de 1,33). Comme pour le son du piano, l'oreille est tolérante.

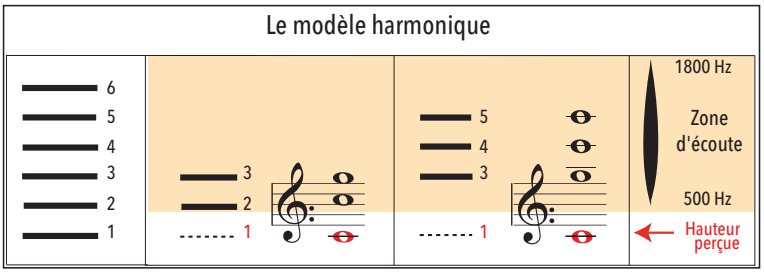


Figure 6.30 Deux composantes harmoniques – ou quasi harmoniques – suffisent pour suggérer un son fondamental dont la fréquence serait le plus grand commun diviseur de ces composantes (voir chapitre 8, page 418). Mais pour que l'effet soit net, il faut que les composantes soient situées dans la zone d'écoute (ou zone dominante), c'est-à-dire entre 500 et 1 800 Hz.

Le modèle harmonique est une forme spectrale prégnante qui nous permet d'interpréter la hauteur des sons inharmoniques de façon économique et efficace (figure 6.30). Il est probable que le codage temporel et les sons différentiels (voir Glossaire) contribuent à renforcer la perception de la hauteur. Le modèle harmonique s'applique à des intervalles approchés, avec une tolérance parfois surprenante. Il prend en compte la zone d'écoute dominante (voir § 3.5.1).

Les exemples abondent. L'un des plus complexes est celui de la cloche d'église[19].

3.4.2. Les notes d'une cloche d'église

L'écoute d'une cloche sonnée à la volée fait mentir l'adage : « qui n'entend qu'une cloche n'entend qu'un son ». Dans notre exemple (Son **6.40**)[20], il est possible à l'oreille de repérer au moins deux ou trois notes, qui diffèrent selon les auditeurs, et de les chanter ou de les siffler.

Son 6.40 (33″)

Sonnerie à la volée d'une cloche d'église

L'analyse sonagraphique du premier coup de cette cloche est donnée figure 6.31. Elle a été volontairement limitée à 2000 Hz pour présenter clairement les principaux partiels qui contribuent à la perception de la hauteur de la cloche. Sur la partie droite de la figure, le spectre instantané effectué à la position du curseur, environ 1,5 s après le choc, permet d'identifier ces partiels avec l'indication de leur note approchée.

19. La perception des notes d'une cloche a fait l'objet de nombreuses recherches aux Pays-Bas (voir Schouten & coll. 1965 ; Collectif-Rossing, 1984).
20. Bourdon de l'église du village de Hochfelden (Haut-Rhin).

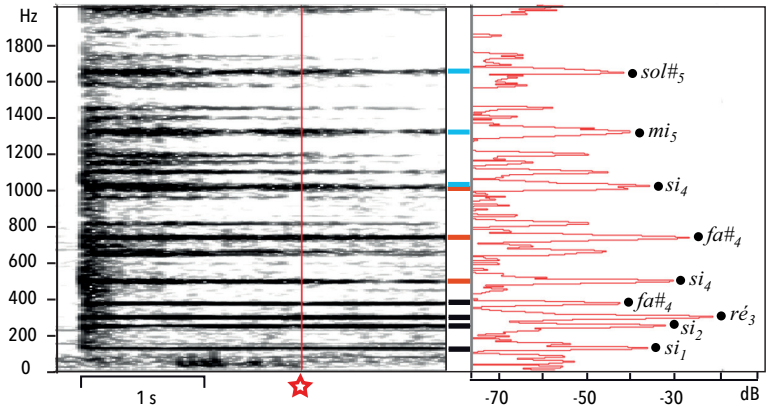


Figure 6.31 Analyse spectrale des premiers partiels de la cloche de Hochfelden. Sonagramme (0-2000Hz ; FFT : 4 096 pts ; Δt = 92 ms) et spectre établi au curseur, montrant les principales composantes avec l'indication des notes approchées.

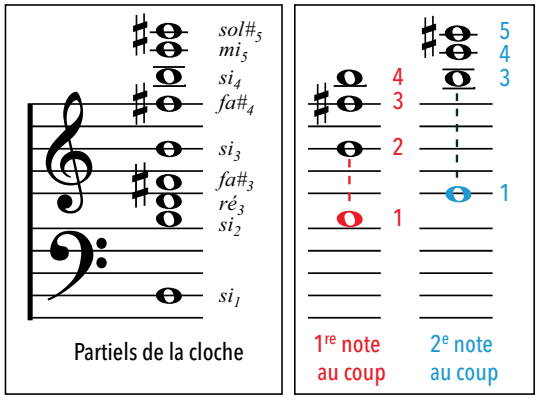


Figure 6.32 Notation musicale approchée des partiels d'une cloche et mise en correspondance des groupes candidats à la perception des deux notes au coup : si_2 et mi_3.

Les modes propres de cette cloche ont été accordés soigneusement par le fondeur. Ils produisent un accord mineur, caractéristique des cloches européennes[21].

1re note au coup

Comme pour la cloche tube, la première *note au coup* d'une cloche de volée, celle qui donne la note fondamentale de la cloche[22], ne correspond pas à une composante réelle du spectre. Elle est le résultat d'une perception associant trois partiels, ici si_3, $fa\#_4$ et si_4 (figure 6.32), dont les fréquences sont voisines des harmoniques 2, 3 et 4 d'un si_2 (en rouge sur la figure). Pourtant, comme les intervalles ne sont pas exacts, la hauteur évoquée par le modèle harmonique peut différer de celle du si_2 présent dans le spectre : le partiel 2 de la cloche. Pour accorder précisément la note de la cloche, le fondeur retouche l'épaisseur aux endroits correspondant aux partiels si_3 et même $fa\#_4$ (voir Van Heuven, E., 1949, p. 94 et 140).

21. L'ajustement des partiels dépend du profil de la cloche. Les cloches d'Orient (Chine, Japon) sonnent très différemment de celle-ci.
22. Le terme fondamental a ici un sens très particulier : c'est la note entendue, demandée par l'usager, la note au coup que doit réaliser le fondeur.

2e note au coup

Les grosses cloches (bourdons) font entendre une deuxième note au coup, à la quarte au-dessus de la première, produite par les partiels d'une autre série de modes vibratoires. Pour la cloche de notre exemple, les partiels concernés sont si_4, mi_5 et $sol\#_5$, bons candidats à la fondamentale mi_3, dont ils pourraient être les harmoniques 3, 4 et 5 (en bleu sur la figure 6.32).

Une expérience très simple permet de vérifier que la quarte entendue provient bien de la présence des partiels aigus (Son **6.41**).

Nous avons reproduit le son de la cloche par synthèse additive. Ensuite, nous avons transposé deux des fréquences de cette série : mi_5 vers $ré\#_5$ et $sol\#_5$ vers $fa\#_5$. À l'audition de la cloche A (Son **6.41a**), modifiée (figure 6.33), seule la première note au coup, si_3, est présente. Dès qu'on rétablit les fréquences originales de la cloche (Son **6.41b**), la note de quarte réapparaît. C'est donc bien le groupe de partiels si_4, mi_5 et $sol\#_5$ qui, sur le principe du modèle harmonique, évoque un mi_3 à la quarte supérieure du si_2.

Son 6.41a (5″)

Cloche A

Son 6.41b (5″)

Cloche B

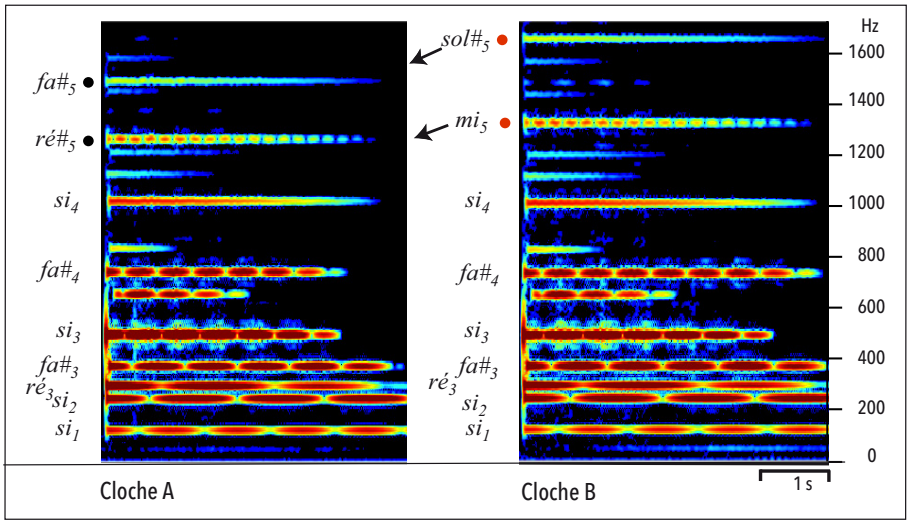


Figure 6.33 Par synthèse additive, on peut simuler le son d'une cloche et translater les fréquences des partiels. Dans la cloche A, la deuxième note au coup a disparu (Son 6.41a).

L'effet de la deuxième note au coup, en quarte, ne se produit pas sur les petites cloches, car les partiels en cause sont à des fréquences beaucoup trop aiguës pour participer à la hauteur perçue. On peut le montrer en transposant le son normal de la cloche (Son **6.42a**) à l'octave supérieure (Son **6.42b**). Non seulement la note de quarte disparaît, mais la tierce mineure, dont la fréquence est maintenant à 600 Hz, devient très saillante. La cloche de Hochfelden, dont les fréquences sont ainsi transposées d'une octave, change de hauteur et de sonorité : c'est une nouvelle cloche.

Voir page 284

Son 6.42a (10″)

Cloche normale

Son 6.42b (11″)

La même cloche transposée à l'octave supérieure

Outre les partiels que nous venons d'étudier, le spectre d'une cloche comporte plusieurs fréquences intermédiaires, car chaque partiel est en fait constitué d'un groupe (un faisceau) de fréquences voisines qui interfèrent et produisent localement de nombreux battements. Ces battements visibles sur les figures 6.31 et 6.33 confèrent une vie extraordinaire au son des cloches, de même que les nombreuses interférences dues aux variations de rayonnement qui se produisent lorsque la cloche se balance au cours d'une sonnerie à la volée.

3.4.3. La note de la timbale

Attention : pour cet exemple, il est très important de disposer d'enceintes restituant bien les basses fréquences ou d'un casque d'excellente qualité.

Son 6.43 (11")

La note de la timbale est un fa, mais dans quelle octave ?

La note d'une timbale s'apprécie assez exactement. Celle de notre exemple (Son **6.43**) est un *fa*, mais l'octave dans laquelle le situe un auditeur est ambiguë.

L'analyse sonagraphique de la figure 6.34 montre, immédiatement après la frappe, un spectre complexe d'où se dégagent pendant l'extinction trois composantes principales[23] – P1, P2, P3 – qui durent 10 à 20 s. La note d'une timbale est d'autant plus juste que ces trois partiels sont bien accordés dans les rapports successifs de la quinte 3/2 pour P2/P1, et de la quarte 4/3 pour P3/P2, ce qui relève de l'art du facteur. Les musiciens savent qu'il se produit une ambiguïté d'octave. Elle provient de deux séries de composantes agissant aussi sur la base du modèle harmonique, l'une dans la partie grave, l'autre dans la partie aiguë du spectre.

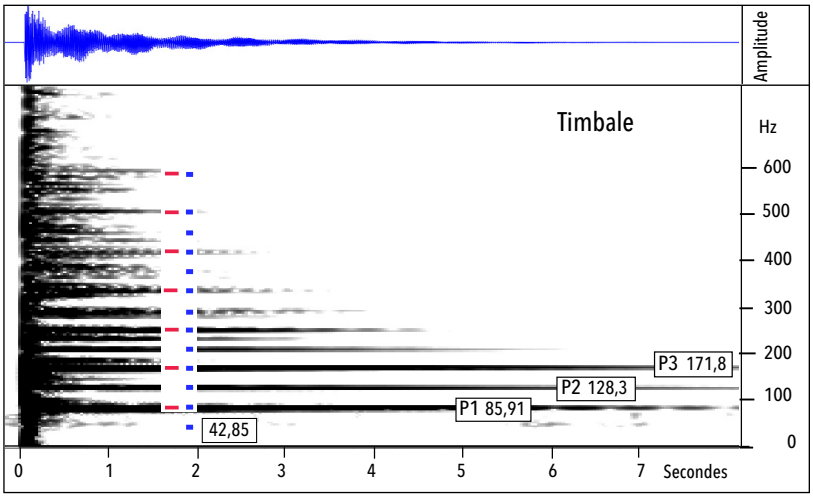


Figure 6.34 Sonagramme d'un coup de timbale. Le réseau harmonique des traits rouges correspond à la note fa_1 ; celui des points bleus correspond à la note fa_0. La note de la timbale est un *fa*, avec une ambiguïté d'octave. Dans les cadres : mesure de la fréquence des partiels en hertz. (Les partiels ont été filtrés puis mesurés avec Praat.)

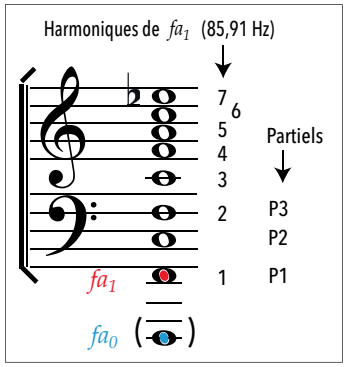


Figure 6.35 Notation musicale des partiels de la timbale montrant l'ambiguïté d'octave du fondamental.

Ainsi on peut entendre (voir la notation musicale figure 6.35) :

- soit un fondamental commun aux trois premiers partiels de l'instrument dont la fréquence, indiquée sur la figure 6.34, est fa_0 = 42,85 Hz (spectre de points bleus) ; l'instrument ne rayonne pas d'énergie à la fréquence du fondamental ;
- soit le fondamental d'un spectre quasi harmonique, tracé en rouge sur la figure, correspondant cette fois à fa_1 = 85,91 Hz.

3.5. La hauteur des sons apériodiques successifs

Dans les pages précédentes, nous avons évoqué l'ambiguïté, la tolérance et les différences individuelles observables à l'écoute de la hauteur des sons apériodiques. Autant d'indices qui montrent que la hauteur perçue de cette catégorie de sons n'est pas établie avec certitude. Elle résulte d'un compromis subtil entre des données acoustiques – les fréquences des composantes, leurs intensités, la tessiture

23. Voir Chaigne & Kergomard, *op. cit.*, page 660, note de bas de page ; voir François, *op. cit.*, p. 86.

spectrale dans laquelle elles se situent – et l'interprétation qu'en donnent les auditeurs selon leurs particularités auditives et leurs habitudes d'écoute. La situation est encore plus complexe lorsque plusieurs sons se succèdent. De fait l'écoute s'inscrit toujours dans un flux temporel, par rapport au son que l'on vient d'entendre et dans l'attente de celui que nous pressentons. L'idée même que l'on peut analyser l'écoute d'un son complexe isolé est presque une utopie puisque le simple fait de le répéter induit fréquemment des changements d'interprétation perceptive. De nouvelles stratégies d'estimation des hauteurs sont à prendre en compte pour l'étude des sons successifs.

3.5.1. La zone d'existence des fondamentaux et et la zone d'écoute dominante

L'expérience de transposition de la cloche (Son **6.42**) suggère l'idée que la hauteur produite par le modèle harmonique ne fonctionne pas sur toute l'étendue du spectre, comme nous l'avons évoqué dans le chapitre 3, page 116. En voici une nouvelle illustration.

Son 6.44 (7")

Deux sons successifs de cloche tube

Écoutons deux cloches tubes l'une après l'autre (Son **6.44**). La première, déjà entendue, sonne do_3. Quel intervalle fait-elle avec la deuxième ? Pour la majorité des auditeurs, la deuxième cloche sonne une tierce majeure plus haut (mi_3) alors qu'elle est en réalité accordée à l'octave supérieure, sur un do_4 ! Que se passe-t-il ?

Comparons sur la figure 6.36 l'analyse spectrale de ces deux tubes. Les notes des partiels les plus intenses sont indiquées sur le sonagramme et transcrites sur une portée musicale (voir figure 6.37). Nous venons de voir (page 256) que la hauteur de la cloche do_3 reposait sur trois partiels : do_4, sol_4 et do_5. L'explication est plus compliquée pour la cloche do_4. Les partiels do_5, sol_5, do_6, qui devraient suggérer une fondamentale à do_4, sont situés beaucoup plus haut en fréquence, entre 1 000 et 2 000 Hz.

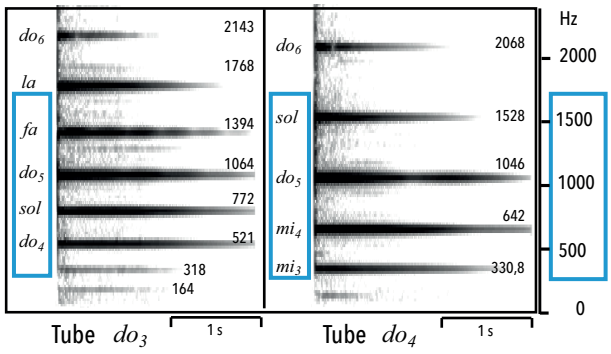


Figure 6.36 Analyse de deux cloches tubes. Les partiels pris en compte pour la perception de la hauteur sont encadrés en bleu. Leur traduction en notes musicales est approchée. En partant des fréquences indiquées on peut estimer la justesse des partiels.

Or plusieurs expériences ont montré[24] (Ritsma, 1962 ; Houtsma,1995) qu'il existait une limite supérieure en fréquence à la perception d'une fondamentale évoquée. D'une part, nous avons tendance à rechercher le fondamental évoqué dans une zone de fréquence comprise entre 150 Hz ($ré_2$) et 440 Hz (la_3), dite zone d'existence. D'autre part, les composantes qui sont de bons candidats à une fondamentale évoquée doivent être situés dans la zone d'écoute dominante qui s'étend environ de 500 à 1800 Hz.

24. Voir chapitre 3, § 2.9 : hauteur des sons complexes et zone de dominance.

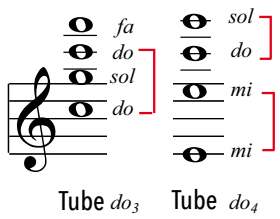


Figure 6.37 Notation musicale approchée des partiels des tubes do_3 et do_4.

Dans le cas de la cloche do_4, le partiel do_6 est très aigu (2000 Hz) et par ailleurs deux composantes intenses sont dans un rapport de fréquences voisin de l'octave (mi_3 et mi_4), ce qui est perceptivement très prégnant (figure 6.37). Finalement, deux hauteurs fondamentales, mi_3 ou do_4, sont en compétition et l'auditeur pourra entendre tantôt l'une, tantôt l'autre, selon le contexte temporel (voir le § 3.5.2). Ces ambiguïtés expliquent que les instruments de musique utilisant des tubes percutés ont une étendue limitée dépassant rarement deux octaves.

On remarque, d'une part, que la « zone d'existence des fondamentaux » correspond pratiquement à celle des fondamentaux de la voix parlée adulte et que, d'autre part, la « zone d'écoute dominante » des partiels contribuant à évoquer un fondamental correspond à celle de l'évolution du deuxième formant vocalique, qui est aussi celle du chant harmonique de style sygyt. Il s'agit bien d'une zone d'écoute privilégiée dans laquelle les aptitudes développées pour le décodage de la parole sont ici dédiées à une écoute raffinée et subtile de la hauteur spectrale.

3.5.2. La règle du plus court trajet spectral

Le paradoxe du son qui refuse d'octavier

L'exemple suivant synthétisé à l'ordinateur par Jean-Claude Risset (1978) est un son paradoxal puisqu'il est à la fois entretenu et totalement apériodique. Le Son **6.45** propose une succession de deux sons complexes. Lorsqu'on interroge les auditeurs sur l'intervalle qu'ils perçoivent entre ces deux sons la plupart disent entendre un intervalle descendant, d'environ un demi-ton. Quelques-uns avancent que le deuxième son paraît plus aigu, mais sans pouvoir préciser l'intervalle.

Son 6.45 (9'')

Un son inharmonique paradoxal

En réalité, le deuxième son est une simple transposition du premier : toutes les fréquences sont doublées. S'il s'agissait d'un son périodique, l'action aurait pour conséquence de produire un saut à l'octave supérieure. Dans le cas présent, les composantes du premier son, dont la hauteur perçue n'est qu'approximative, sont disposées de façon très particulière. Les fréquences de ces composantes sont (du grave à l'aigu) : 29,52 ; 60,8 ; 125,3 ; 258, etc. En doublant ces fréquences, on obtient respectivement : 59,04 (qui est plus bas que 60,8), 121,6 (qui est plus bas que 125,3), 250,6 (qui est plus bas que 258) et ainsi de suite. La logique d'écoute consiste à prendre le chemin le plus court en comparant les composantes faisant le plus faible intervalle (un demi-ton descendant plutôt qu'une 7e majeure ascendante).

Son 6.46 (49'')

Cloche tube : gamme chromatique : 13 sons, puis le 13e et le 1er

La règle du plus court trajet mélodique peut aussi avoir un effet stabilisateur. Revenons un instant sur les cloches tubes pour écouter une gamme chromatique partant de do_3 (Son **6.46**). Parvenu au treizième son, nous entendons cette fois clairement un do_4 pour la cloche du Son **6.44** que nous entendions comme un mi_3 précédemment. De proche en proche, notre écoute est restée fixée sur les trois partiels qui produisent le fondamental évoqué sur un modèle harmonique. Cependant en fin de séquence, la juxtaposition des sons 1 et 13 rompt cette similitude et l'intervalle d'octave disparaît, pour la plupart des auditeurs, au profit de l'intervalle de tierce, plus prégnant dans ce contexte.

3.5.3. L'émergence par saillance temporelle, ou comment jouer deux notes avec une cloche

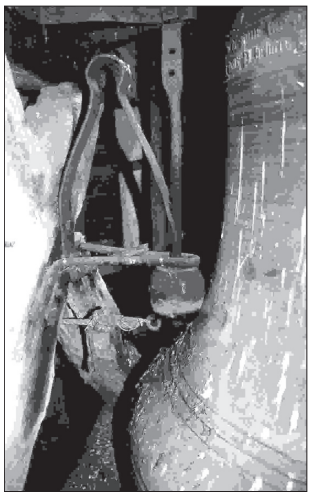


Figure 6.38 Marteau roulant percutant la cloche de l'extérieur, au niveau de la ceinture.

Cloche de Fains (Eure) ; cliché Régis Singer.

Lors d'une mission d'enregistrement en Normandie, quelle ne fut pas notre surprise d'entendre une sonnerie produisant deux notes distinctes successives, alors qu'on nous avait assuré que l'église ne possédait qu'une cloche[25] (Son **6.47a**) ! Un dispositif ingénieux permettait de faire alterner la frappe habituelle du battant sur le bord intérieur de la cloche, avec la frappe d'un marteau extérieur dit marteau roulant, qui excite la cloche dans sa partie cintrée vers le milieu du profil (figure 6.38). Il est très difficile de repérer à première vue les deux notes de cette cloche sur un sonagramme, car les partiels produits par le battant intérieur sont très intenses et surtout parce qu'ils persistent longtemps après la frappe (figure 6.39). Nous savons (chapitre 1, page 16) que le point de frappe opère une sélection dans la série des modes. Le seul fait que le marteau roulant change l'intensité relative des partiels prend, dans la succession temporelle, un poids perceptif qui explique l'émergence de la tierce mineure (*sol*), car « c'est ce qui change qui attire l'attention ».

Son 6.47a (32")

Sonnerie du glas avec la cloche de Ry
(voir page 284)

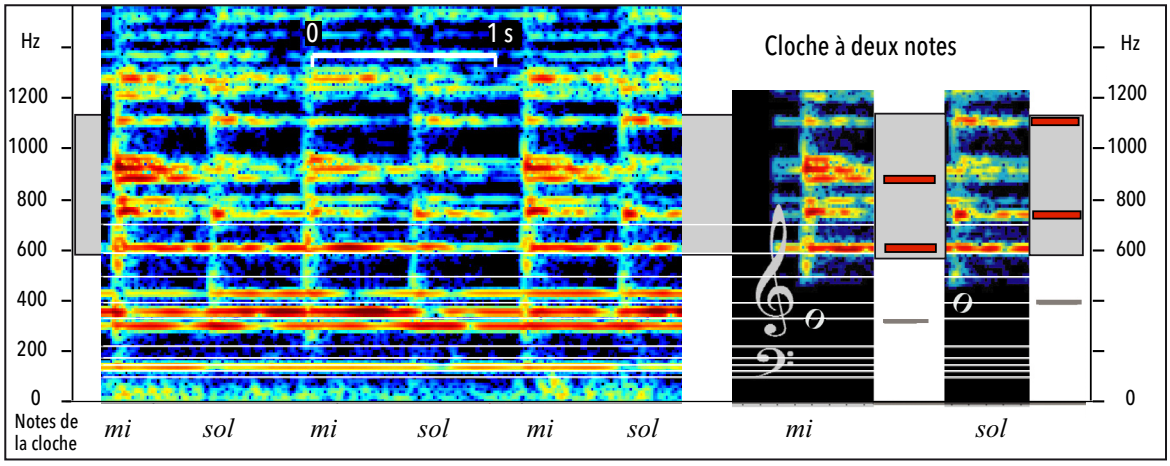


Figure 6.39 Analyse sonagraphique de la cloche de Ry sonnant deux notes différentes. À gauche, spectre de 0 à 2 000 Hz ; à droite, zone filtrée entre 500 et 1200 Hz, montrant les deux partiels responsables de la perception des deux notes, *mi* et *sol*. Ces partiels, situés dans la zone d'écoute dominante, sont « comme » les harmoniques 2 et 3 de la note entendue.

Après plusieurs essais de filtrage et d'écoute, nous avons repéré la zone de fréquence qui contribue à la perception de hauteur de la deuxième note de la cloche. Elle se situe une fois encore dans la zone dominante mentionnée plus haut : entre 600 et 1500 Hz (figure 6.39 à droite et Son **6.47b**).

Son 6.47b (5")

Filtrage du son de la cloche entre 500 et 1 200 Hz
(voir page 284)

25. Il s'agit de la commune de Ry (Seine-Maritime). D'autres marteaux roulants existent dans le département de l'Eure : à Isneauville, Saint-Pierre-de-Bailleul et Saint-Marcel qui ont été documentés par Régis Singer, campanologue auprès du ministère de la Culture. Cet astucieux dispositif semble circonscrit, en France, à une région très réduite.

D'anciennes cloches chinoises, de forme ovale, produisent aussi deux notes. La fouille de la tombe du marquis Yi de Zeng en Chine a révélé un carillon de 64 cloches accordées, datant de 2000 ans avant notre ère, sur lesquelles deux points de frappe différents sont indiqués. Ils correspondent aussi à un intervalle voisin de la tierce, cependant la géométrie de ces instruments est fort différente de la cloche étudiée ici (voir Shen, 1987 ; Rossing, 1990). La règle du plus court trajet spectral qui permet d'expliquer l'émergence de mélodies de sons inharmoniques s'applique aussi aux sons fortement amortis, pourvu que soient repérables des maxima spectraux.

3.5.4. Les percussions fortement amorties

La grande famille des instruments à percussion se partage entre instruments accordables et instruments de hauteur indéterminée[26], ces derniers ayant généralement un fort amortissement de la vibration.

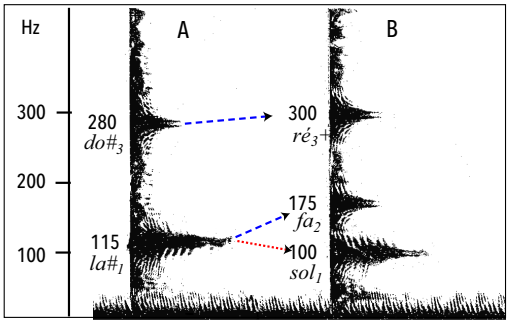


Figure 6.40 Analyse spectrale de chocs sur deux tables d'harmonie différentes et figuration de trajets mélodiques possibles selon les auditeurs.

Son 6.48 (11")

Percussions sur deux tables d'harmonie de luth (répété deux fois)

Le Son **6.48**, produit par la percussion de deux planchettes de bois sélectionnées pour être des tables d'harmonie est un exemple typique de hauteur ambiguë. Le luthier recherche celle qui sonne le plus clair, le plus haut. Lorsqu'on fait entendre successivement les deux sons A et B à un groupe d'auditeurs, la moitié déclare que A est plus haut que B et l'autre moitié est de l'avis contraire. Bien sûr les deux réponses sont bonnes et tout à fait explicables lorsqu'on examine l'analyse spectrale des sons (figure 6.40). L'impression de montée ou de descente mélodique que suggère la succession A B peut d'ailleurs s'inverser d'une écoute à l'autre : selon la zone de tessiture choisie. Pour réduire l'ambiguïté, les instruments de percussion à membrane sont généralement munis d'un fût de résonance dont l'accordage permet de focaliser la hauteur. L'expérience de filtrage passe-haut par bandes successives, de l'aigu vers le grave, en montre l'importance. (Son **6.49**).

Son 6.49 (1'52)

Jeu sur trois tom-tom

Filtrages successifs de l'aigu au grave *(voir page 285)*

La perception de hauteur dont nous parlons à propos des sons inharmoniques, purement spectrale, fait partie intégrante de ce qu'on nomme de façon générale « le timbre » et devient un indice d'identification, un repère spectral qui permet de suivre la séquence rythmique d'un instrument donné dans un ensemble. Les différents tambours d'une polyrythmie africaine (ou les différents sons d'un tambour de hauteur variable) sont suffisamment distincts pour être discriminés et leur nombre dépasse rarement quatre ou cinq. La catégorisation de hauteurs spectrales ainsi mise en œuvre n'est sans doute pas éloignée de celle que nous déployons pour la discrimination des voyelles de la langue. On remarquera d'ailleurs que maints percussionnistes « parlent » leurs rythmes et que, symétriquement, les tambours peuvent transmettre des messages parlés dans les langues à ton, celles dont le sens des mots dépend d'une intonation de type contour mélodique.

26. Il serait plus exact de dire hauteur imprécise ou ambiguë.

4. Les sons périodiques modulés : vibratos et trilles

4.1. Introduction : les instabilités de fréquence des sons réels

Jusqu'à présent nous avons analysé la perception de hauteur de toutes sortes de sons dont la fréquence était considérée comme stable. Du point de vue perceptif, un son stable est un son qui paraît dépourvu de variations à l'écoute. Pour y voir de plus près, nous avons demandé à plusieurs musiciens de produire des sons « droits », sans vibrato.

Le Son **6.50** fait entendre trois sons la_3, joués successivement à la flûte à bec et au violon puis chanté, qui sont effectivement bien stables. Or nous sommes plutôt surpris de découvrir sur l'analyse des fluctuations de fréquence dont nous n'avions pas conscience (voir figure 6.41).

Son 6.50 (12")

Sons « droits », note la_3

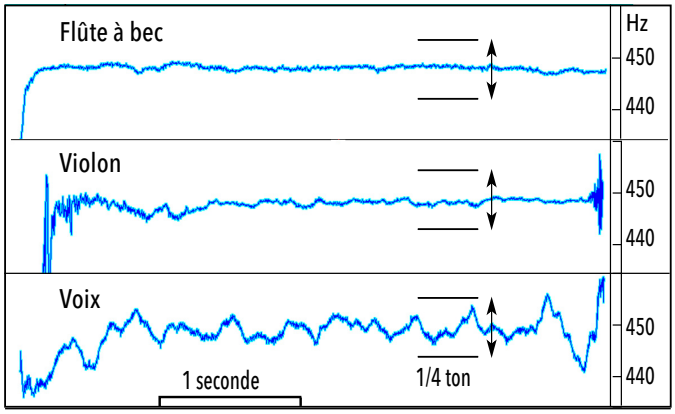


Figure 6.41 La stabilité en fréquence d'un son joué le plus « droit » possible à la flûte à bec ou au violon est toute relative : on observe des irrégularités, mais qui sont très faibles lorsqu'on les compare à celles de la voix humaine.

Il est intéressant de remarquer que les fluctuations, très faibles pour les deux instruments (de l'ordre de 4/1 000 de Hz), sont trois fois plus notables à la voix (13/1 000), soit 3 % de la fréquence, ce qui est de l'ordre du 1/8 de ton.

Il est plus facile de stabiliser la production d'un son de hauteur donnée lorsqu'existe un feedback entre l'excitation (archet, souffle) et la structure vibrante (corde ou tuyau). En ce qui concerne la voix chantée, la stabilité n'est assurée que par l'équilibre des contrôles musculaires du souffle et de l'appareil phonatoire.

Si le son vocal de cet exemple paraît droit à nos oreilles, c'est parce que nous avons appris à ignorer les inévitables fluctuations d'une voix humaine de la catégorie « voix chantée ». On peut avancer l'hypothèse que le développement du vibrato de fréquence sur les notes tenues, dont on sait l'importance dans la musique occidentale, serait une façon de masquer ces variations, en imposant à la fréquence fondamentale une modulation périodique dont l'échelle temporelle est beaucoup plus grande.

4.2. Le vibrato musical : une modulation complexe

La perception des modulations d'amplitude et de fréquence de sons purs est exposée au § 3.2 du chapitre 3 (système auditif). Nous analyserons ici celle des techniques musicales de modulation, plus connues sous le nom de vibrato, qui sont en usage dans le chant et dans le jeu des instruments.

Le vibrato est une modulation lente du son, inférieure à huit modulations par seconde. Selon les instruments, cette modulation s'applique préférentiellement à l'amplitude ou à la fréquence. Dans tous les cas, le vibrato s'accompagne aussi d'une modulation du spectre.

4.2.1. Le vibrato d'amplitude

Quelques instruments ont pour moyen d'expression un vibrato d'amplitude, dû à un système mécanique ou à une registration particulière. Notre sensibilité aux variations d'amplitude est optimale à 4 Hz (voir chapitre 3, page 121). Dans le jeu du vibraphone, le vibrato d'amplitude est produit par la variation périodique du couplage (par résonance) d'un tuyau placé sous la lame et accordé à la fréquence du premier mode de celle-ci. À l'orgue, il existe un vibrato d'amplitude produit par le registre de tremblant doux qui provoque une modulation du débit d'air dans le conduit d'alimentation. En revanche, l'ondulation du jeu de voix céleste provient du fait qu'il est accordé 4 Hz plus haut que les autres jeux de l'instrument. Lorsqu'on tire simultanément un jeu de flûte et celui de voix céleste, le vibrato d'amplitude ainsi obtenu est dû aux battements qui se produisent entre deux tuyaux désaccordés, et non à la modulation d'amplitude d'une seule source. Il en est de même du registre dit de musette de l'accordéon.

Son 6.51 (31")

Voix chantée : vibrato d'amplitude, puis de fréquence

Exceptionnel en technique vocale, le vibrato d'amplitude produit un effet expressif très particulier. Le Son **6.51** donne à entendre un magnifique exemple enregistré en 1940 au Liban. Le chanteur amorce un vibrato d'amplitude très visible sur la forme d'onde de la figure 6.42, qui commence lentement (2 modulations par seconde), va en s'accélérant et se transforme au bout de 12 secondes en vibrato de fréquence. Particulièrement large (271 cents), ce dernier s'accompagne également d'un vibrato d'amplitude.

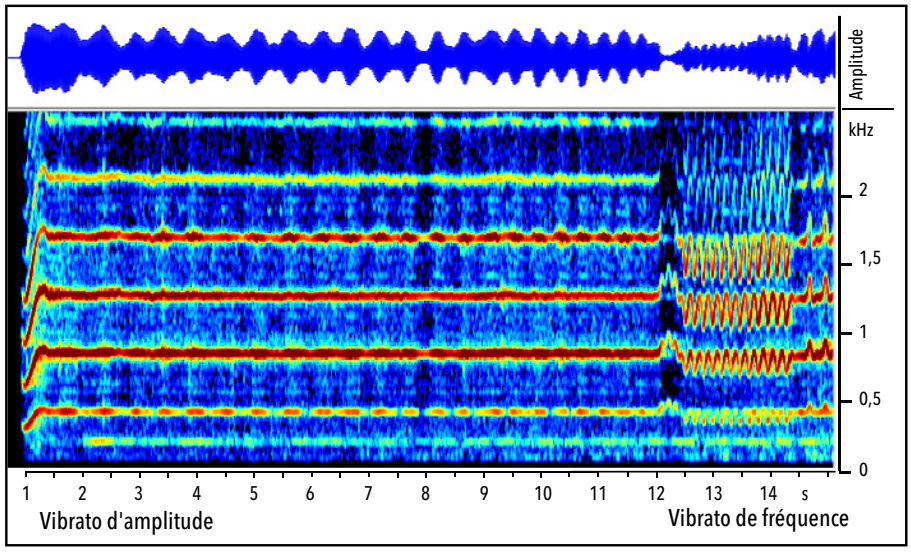


Figure 6.42 Analyse d'un chant syrio-libanais avec, successivement, un vibrato d'amplitude et un vibrato de fréquence.

De façon générale, les instruments à vent peuvent produire d'importantes modulations d'amplitude résultant des variations de débit aérien, toujours associées à des modulations de fréquence et de spectre. Le Son **6.52a**, un *do*$_3$ de flûte, est analysé figure 6.43. Il est animé d'une modulation d'amplitude, bien visible sur la courbe supérieure, dont la fréquence est de 5,5 par seconde. Dans le plan spectral, on constate d'importantes variations synchrones du spectre harmonique, associées à une faible modulation de la fréquence de jeu. Cet exemple est représentatif d'un style de jeu, et en particulier d'une technique de souffle qui n'est pas du goût de tous les flûtistes. D'autres préfèrent un vibrato où les variations de fréquence sont plus marquées.

Son 6.52a (4")

Flûte traversière, vibrato (do$_3$).

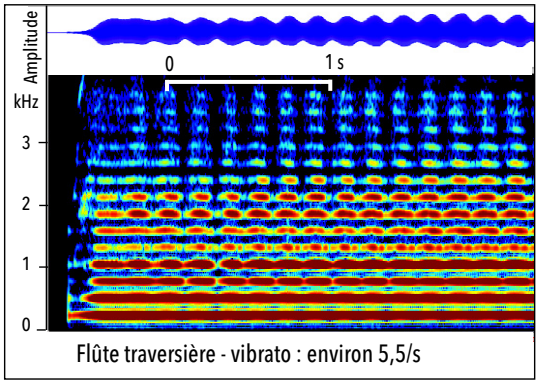


Figure 6.43 Exemple d'un vibrato d'amplitude à la flûte traversière. Note *do*$_3$; FFT : 1 024 pts.

Son 6.52b (4")

Violon, vibrato (do$_3$).

4.2.2. Le vibrato de fréquence

Le vibrato de fréquence est caractéristique des instruments à archet (quatuor), du trombone à coulisse et principalement des chanteurs. Notre exemple de vibrato du violon combine des variations de fréquence (6 Hz), d'amplitude et de spectre (Son **6.52b** et figure 6.44). Le vibrato du chanteur (Son **6.52c** et figure 6.45) est avant tout un vibrato de fréquence, plus ample et plus irrégulier (pour cet interprète) que celui du violon.

Son 6.52c (4")

Voix chantée vibrato (do$_3$).

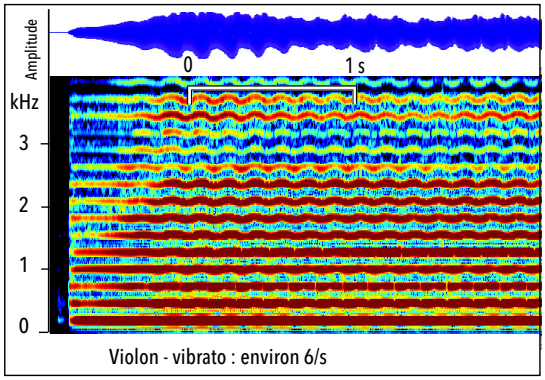


Figure 6.44 Son vibré joué au violon, note *do*$_3$. Le vibrato affecte la fréquence, l'amplitude et le spectre ; FFT : 1 024 pts.

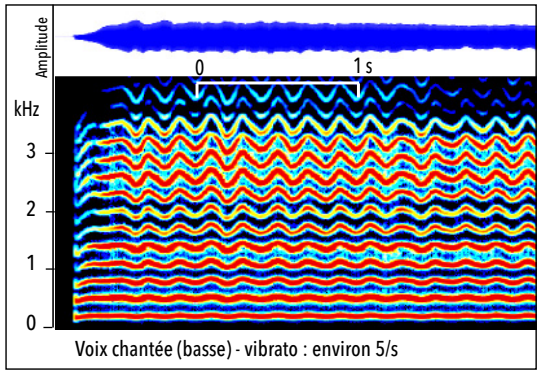


Figure 6.45 Son chanté avec vibrato, note *do*$_3$. Notez l'amplitude des variations de la modulation en fréquence ; FFT : 1 024 pts.

Nous avons choisi de comparer ces trois exemples sur la même note, mais il faut garder présent à l'esprit le fait que le vibrato varie énormément d'une note à l'autre.

La cadence de modulation recherchée pour le vibrato de fréquence n'est pas celle du maximum de sensibilité au suivi de la variation, ce qui produirait l'impression désagréable d'un « pleurage », mais celle du seuil de fusion. L'optimum avoisine 6 Hz, avec une marge de variation (de 5 à 7,5 Hz) qui dépend de la tessiture, du tempo et principalement du goût des musiciens[27]. L'appréciation du vibrato de fréquence s'acquiert, surtout s'il est ample comme en chant lyrique ; certains auditeurs y demeurent réticents malgré leurs efforts.

La production d'un vibrato en fréquence résulte d'un mouvement de la main (violon), d'une oscillation du larynx (voix) ou des variations de pression de l'air fourni par le musicien (flûte, hautbois). La plage des valeurs observées est donc aussi à mettre en relation avec les fréquences d'oscillation des systèmes physiologiques moteurs mis en jeu dans sa réalisation.

4.3. Le vibrato de fréquence et la perception de la hauteur

4.3.1. Les incidences de la largeur du vibrato

La *largeur* du vibrato, ou intervalle d'excursion en fréquence, varie également beaucoup selon les instruments. Le vibrato le plus ample est celui des chanteurs et le plus faible est celui de la clarinette dans un style de jeu classique.

En musique, la largeur de modulation est définie par le rapport de l'intervalle de modulation Δf (mesuré en Hz) à la fréquence moyenne fm[28]. Ce rapport, $\Delta f/fm$ (figure 6.46), ne dépend pas du rang de l'harmonique. Les valeurs courantes de $\Delta f/fm$ sont 0,03 (soit 3 % de la fréquence) pour un vibrato de violon et 0,06 à 0,12 pour un chanteur. Ces trois valeurs correspondent respectivement à 1/4, 1/2 et 1 ton tempéré.

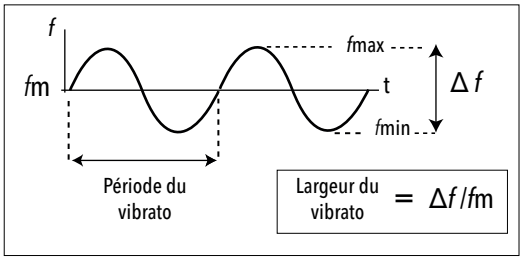


Figure 6.46 Schéma du vibrato de fréquence.

Son 6.53 (25")

Différentes largeurs de vibrato : son de synthèse *(voir page 285)*

Le Son **6.53** réalisé en synthèse permet d'apprécier comment, pour chacun de nous, la hauteur perçue varie avec la largeur du vibrato (se reporter à la liste des sons pour le détail). On remarquera dans cet exemple que, pour la largeur maximale du vibrato (un ton), la hauteur perçue de la note vibrée n'est plus comparable à celle du son non vibré. Selon les auditeurs, elle paraît soit plus haute, soit plus basse !

27. Il est frappant de constater que les chanteurs lyriques de culture latine (Italie) apprécient des fréquences de vibrato plus rapides que ceux de culture anglo-saxonne.
28. Dans certains ouvrages, Δf est défini comme l'écart par rapport à la fréquence moyenne fm ; cette pratique vaut pour une modulation symétrique de part et d'autre de la fréquence moyenne, ce qui est rarement le cas dans la réalité musicale. L'intervalle perçu correspond à l'excursion totale, pour laquelle il n'existe pas de terme bien défini : largeur, excursion ou amplitude du vibrato sont indifféremment employés.

Ce n'est donc plus la fréquence moyenne que nous percevons mais plutôt la limite supérieure ou inférieure de l'excursion en fréquence. Nous retrouverons ce mode de perception dans l'étude du trille vocal, page 272.

4.3.2. Les tolérances de justesse induites par le vibrato de fréquence

Les déplacements de la hauteur perçue, vécus dans l'expérience du son **6.53** indiquent que, dès qu'un son est vibré, l'auditeur devient beaucoup plus tolérant pour apprécier la justesse des intervalles musicaux. L'écoute du Son **6.54** est surprenante. La mélodie de trois notes, jouée avec vibrato sur un instrument électronique, est plutôt bien acceptée. Sans vibrato, l'écart de justesse n'est plus du tout supportable. On en conclut que le vibrato permet à l'auditeur d'interpréter la hauteur moyenne, à la baisse ou à la hausse, selon le contexte. Écoutez aussi le Son **6.55** : quel est l'intervalle entre les deux sons chantés ?

Son 6.54 (17″)

Vibrato et tolérance de justesse (voir page 285)

Son 6.55 (9″)

Repérez l'intervalle chanté

4.3.3. Le vibrato de fréquence dans les transitions mélodiques

Vibrato et tempo

Dans la pratique musicale, le vibrato affecte des sons de durées variables, liés ou détachés. Les interprètes s'arrangent, dans le jeu legato, pour effectuer un nombre entier d'oscillations pendant la durée d'une note. Que se passe-t-il lorsque le nombre de notes par seconde avoisine la fréquence du vibrato ? Le Son **6.56** permet de comparer la réalisation d'une gamme diatonique ascendante-descendante de do_3 à sol_4, chantée puis jouée au piano. À l'écoute, il s'agit bien de la même mélodie.

Sur l'analyse spectrographique du piano (figure 6.47), il est aisé de repérer les différentes notes qui sont autant de traits horizontaux séparés par le bruit de la percussion. L'interprétation mélodique de la gamme chantée s'avère plus difficile.

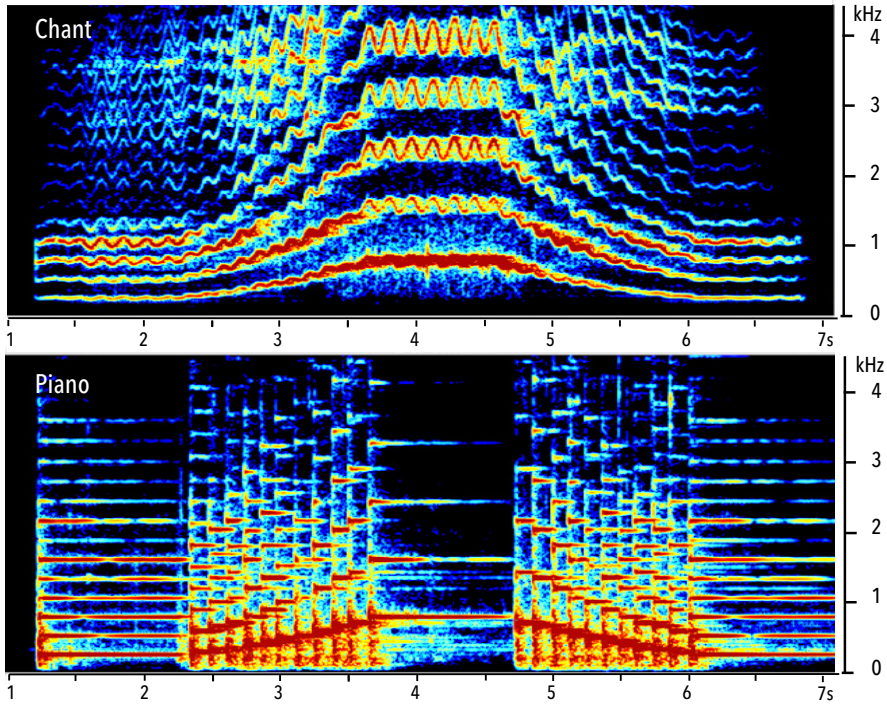


Figure 6.47 Gamme diatonique rapide chantée, puis jouée au piano. Son 6.56.

Son 6.56 (15″)

Gamme diatonique chantée (soprano) puis jouée au piano

Les deux notes extrêmes tenues ont plusieurs arches de vibrato (5 pour le *do* grave et 7 pour le *sol* aigu). Sur toutes les autres notes, la chanteuse « passe » par une seule arche (voir figure 6.48). L'enchaînement ne se fait pas au hasard. La transition à la note suivante s'opère dans la partie ascendante du vibrato, pour une note supérieure, et dans la partie descendante pour une note inférieure.

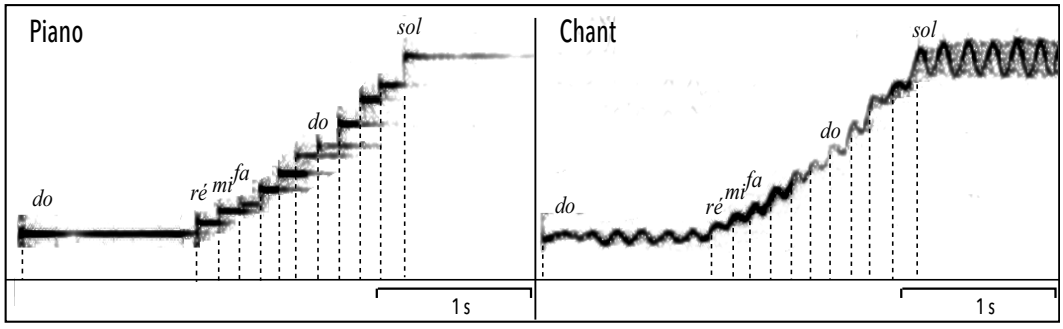


Figure 6.48 Ci-dessus, ligne mélodique de la gamme diatonique do_3 à sol_4, jouée au piano et chantée avec un vibrato de fréquence. Le tracé est celui de l'harmonique 3 dans la partie ascendante de la gamme.

En comparant les traits précis de la fréquence fondamentale de chacune des notes du piano et les larges ondulations de la ligne mélodique vibrée, on ne peut qu'insister sur la part importante que prend l'interprétation cognitive dans la perception de hauteur des sons vibrés. Toutefois, plusieurs expériences de synthèse ont permis de montrer que la marge est réduite et qu'il est tout à fait possible d'ajuster de façon précise des sons vibrés (Castellengo & coll., 1989).

Son 6.57 (30")

Vibrato dans le portando

Lorsque la transition se fait entre deux notes liées séparées par un très grand intervalle, comme dans le Son **6.57** où il est d'une octave + une quinte diminuée (du $ré_3$ au lab_4) la durée de transition peut atteindre des valeurs importantes : ici 400 ms. Un simple raccordement glissé entre les deux notes serait alors perçu comme un portamento peu agréable.

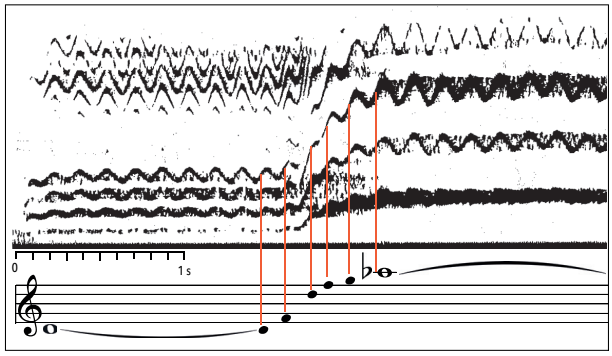


Figure 6.49 Transition vibrée entre deux notes liées, séparées par un grand intervalle (Son 6.57).

Dans notre exemple la chanteuse réalise une transition très élaborée dont nous n'avons pas conscience à l'écoute. L'analyse (figure 6.49) révèle qu'elle effectue un

arpège de trois oscillations intermédiaires sur fa_3, $ré_4$, fa_4, qui sont les notes de l'accord de quinte diminuée. Ces notes intermédiaires ne sons pas perçues à la première audition et, d'ailleurs, ne sont pas écrites sur la partition. Un simple calcul montre que pendant la transition la période des oscillations se réduit : elle passe de 160 ms (vibrato de 6, 25/s) à 133 ms. L'habileté musicale de la chanteuse, qui ne doit pas arrêter l'oscillation de sa voix pendant cette liaison, consiste à effleurer des degrés s'inscrivant dans l'harmonie de l'accord.

4.3.4. La perception de hauteur des notes courtes vibrées

L'exemple le plus célèbre est celui de l'air de la Reine de la nuit de *La Flûte enchantée* de Mozart qu'on peut entendre dans le Son **6.58**. La figure 6.50 montre (ligne B), la transcription mélodique extraite de l'analyse spectrale des huit notes piquées sur la_4.

Son 6.58 (18″)

Les notes « piquées » de la Reine de la nuit

Chaque note est une portion d'arche de vibrato d'une durée moyenne de 80 ms, soit une demi-période de vibrato. Des expériences systématiques d'ajustement de la hauteur perçue pour ces brèves modulations, telles qu'elles se présentent dans l'interprétation chantée, ont montré que la hauteur perçue n'est plus assimilable à la fréquence moyenne (*f*m), mais varie avec le sens ascendant ou descendant de l'inflexion et en particulier avec sa forme terminale[29].

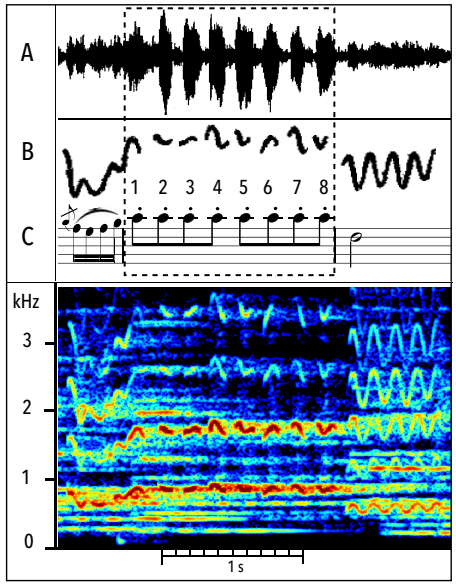


Figure 6.50 Analyse acoustique d'un extrait de l'air de la Reine de la nuit (Mozart) montrant la diversité des réalisations de la note « piquée » répétée 8 fois (note la_4). De haut en bas : A, courbe d'amplitude ; B, tracé de l'harmonique 4 ; C, notation musicale ; sonagramme.

Son 6.59 (4″)

Synthèse des formes a à d de la figure 6.51

Les formes schématiques a et d de la figure 6.51 sont perçues plus haut que les formes b et c (Son **6.59**). Le Son **6.60a** donne une idée du test qui consiste à apparier la hauteur perçue d'une forme avec celles de notes stables de fréquence variable. Lorsque la durée s'allonge, la hauteur perçue de la forme b reste stable (Son **6.60b**).

Son 6.60a (12″)

Test des notes courtes vibrées (voir page 285)

Test 1

Son 6.60b (5″)

Test 2

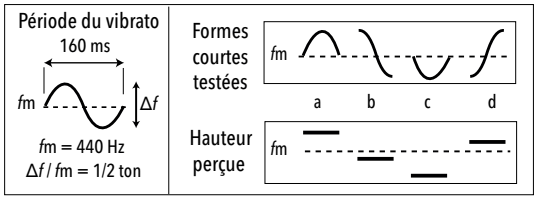


Figure 6.51 Schémas des portions de vibrato (a, b, c, d) qui ont fait l'objet de tests et figuration de la hauteur perçue par rapport à la fréquence moyenne.

29. Voir d'Alessandro C., Castellengo M., *The pitch of short-duration vibrato tones* (1993) et (1994).

4.4. Du vibrato au trille : le rôle du contexte musical

Le cas étrange du trille vocal

Le trille est un ornement mélodique commun à tous les instruments, constitué de l'alternance de deux notes conjointes à intervalle d'un demi-ton ou d'un ton. Il se distingue nettement du vibrato par la netteté des deux degrés en présence. La cadence du trille est approximativement celle du vibrato : autour de six à huit alternances par seconde.

C'est le trille vocal, difficile à bien réaliser, qui va retenir notre attention. Comment un chanteur peut-il effectivement produire l'alternance rapide de deux notes ? Comparons tout d'abord un trille instrumental et un trille vocal.

Son 6.61 (18″)

Trille vocal et trille instrumental

Le Son **6.61** propose successivement un trille vocal et un trille joué en écho à la flûte, une octave plus haut. L'écoute en est très claire. Nous avons effectué l'analyse sonagraphique et extrait l'harmonique 2 de la voix pour le comparer au fondamental de la flûte, afin de positionner les deux tracés en vis-à-vis (figure 6.52).

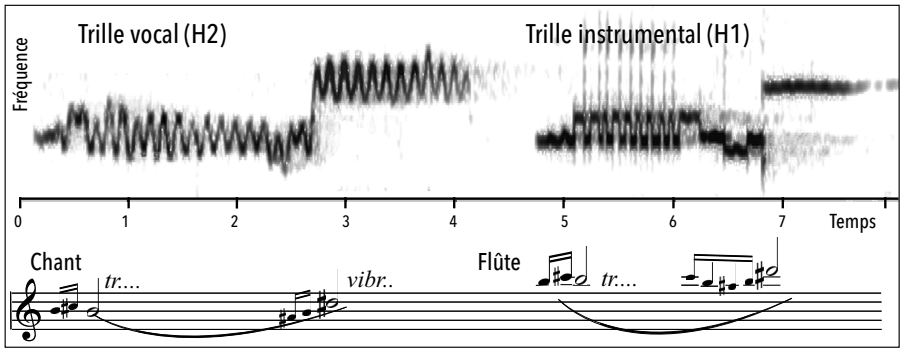


Figure 6.52 Un trille vocal suivi d'un trille joué à la flûte. Les deux trilles étant à une octave de distance dans le Son 6.61, nous avons sélectionné l'harmonique 2 de la voix et l'harmonique 1 de la flûte pour les comparer à la même fréquence moyenne.

Son 6.62 (5″)

Trois sons vibrés (figure 6.53 à gauche)

Son 6.63 (5″)

Trille avec préparation (figure 6.53 à droite)

Son 6.64 (5″)

Extraits (1) et (2) de la figure 6.53

Le tracé du trille de la flûte, dont les deux notes (*si* et *do*#) sont séparées temporellement par le bruit de fermeture des clés du *si* (on en compte 9 par seconde), est clairement lisible sur l'analyse. En revanche, il n'est pas facile de s'y retrouver dans la lecture du trille vocal. La figure montre les ondulations de deux sons vibrés de largeur équivalente, correspondant au si_3 et au $ré\#_4$: comment se fait-il que le premier soit perçu comme un trille et le deuxième comme une note vibrée ? Cette observation nous a suggéré l'idée de réaliser une expérience de synthèse.

Dans le Son **6.62**, la voix que nous entendons chante trois notes vibrées ; dans le Son **6.63**, la voix chante un trille. Or nous sommes bien étonnés d'apprendre que le signal synthétisé pour la dernière note vibrée, $do\#_4$, est le même que celui de la note trillée, ce que l'on peut constater sur l'analyse de la figure 6.53. Il suffit pour s'en convaincre d'écouter les sons (1) et (2) de la figure 6.53, isolés du contexte (Son **6.64**). Comment s'opère le changement de perception entre la hauteur unique d'une note vibrée et l'alternance des deux notes distinctes d'un trille ?

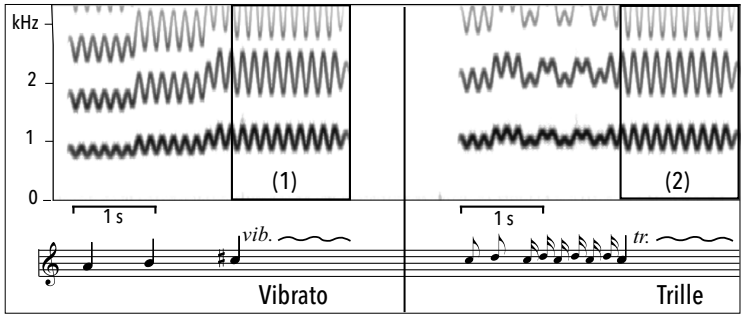


Figure 6.53 Un même signal vibré (encadré), peut être perçu soit comme une seule note avec un vibrato, soit comme un trille de deux notes.

Nous avions déjà remarqué (page 268) que la hauteur perçue d'un son vibré largement pouvait se déplacer, les limites supérieures ou inférieures de l'excursion pouvant notamment devenir de nouveaux repères. C'est exactement ce qui se passe lorsque nous entendons un trille vocal. Le schéma de la figure 6.54 permet de comprendre comment un son vibré largement s'entend, tantôt comme une seule note, tantôt comme les deux notes d'un trille. C'est donc seulement dans un contexte musical donné que le signal prend sens, qu'il est interprété comme vibrato ou comme trille[30].

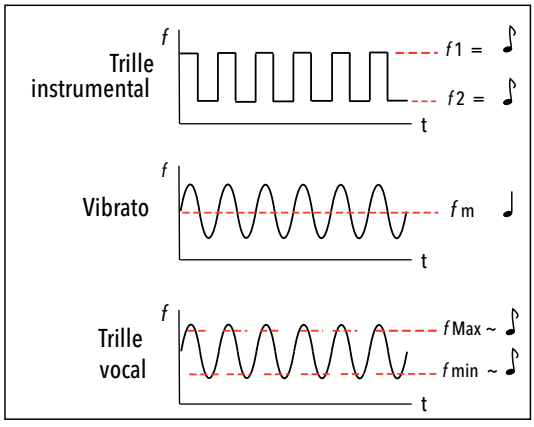


Figure 6.54 Le trille instrumental est dû à l'alternance de deux notes distinctes. Le trille vocal n'est qu'un grand vibrato. Lorsque la largeur de l'excursion a des valeurs comprises entre 0,12 et 0,18, la perception est totalement ambiguë et ne dépend que du contexte musical.

Les exemples sonores suivants confirment le rôle important du contexte musical. Dans le Son **6.65**, l'intervalle exact d'un trille isolé n'est pas facile à appréhender : est-ce un ton ou demi-ton ? Il est difficile de se prononcer. Avec la préparation, tout s'éclaire. Dans le Son **6.66**, les deux notes de préparation du trille et la formule conclusive tonale sont la clé d'écoute musicale du trille vocal.

Son 6.65 (8″)

Deux trilles chantés hors contexte musical

Son 6.66 (25″)

Les mêmes trilles chantés en contexte

Le trille vocal est peut-être la plus belle illusion auditive, musicalement partagée. En la comparant à la parabole du cheveu de la Vierge (voir annexe D) Stéphen de la Madelaine, éminent chanteur et pédagogue écrit :

> *Je me suis cent fois remémoré cette fraude pieuse en écoutant en fort bons lieux des trilles qui escamotaient si habilement la note supérieure, que le public croyait entendre un pianissimo de la plus exquise délicatesse, tandis qu'en réalité l'artiste faisait sautiller sa voix sur une seule note.*

30. Voir Castellengo, M. (1993), *Fusion or separation : from vibrato to vocal trill*, Comptes-rendus du SMAC93 (Stockholm Musical Acoustics Conference) Stockholm.

D'autres chanteurs se bornent à soulever la première note du trille, de manière à obtenir une seconde diminuée, d'une justesse équivoque, et le rapide martellement du trait (quand il y a martellement) produit encore, dans cette circonstance, une illusion qui satisfait beaucoup d'oreilles. Mais le seul trille régulier, tel que le battent madame Damoreau et Ponchard, est celui qui produit nettement et vigoureusement les deux notes qui le composent.

Madelaine, S. (de la), vers 1851, p. 119-121.

Son 6.67 (21")

Trille vocal avec préparation et terminaison

Le Son **6.67** associé à l'analyse de la figure 6.55 offre l'exemple d'une très belle réalisation de trille vocal, digne de la « Damoreau », par la chanteuse Joan Sutherland.

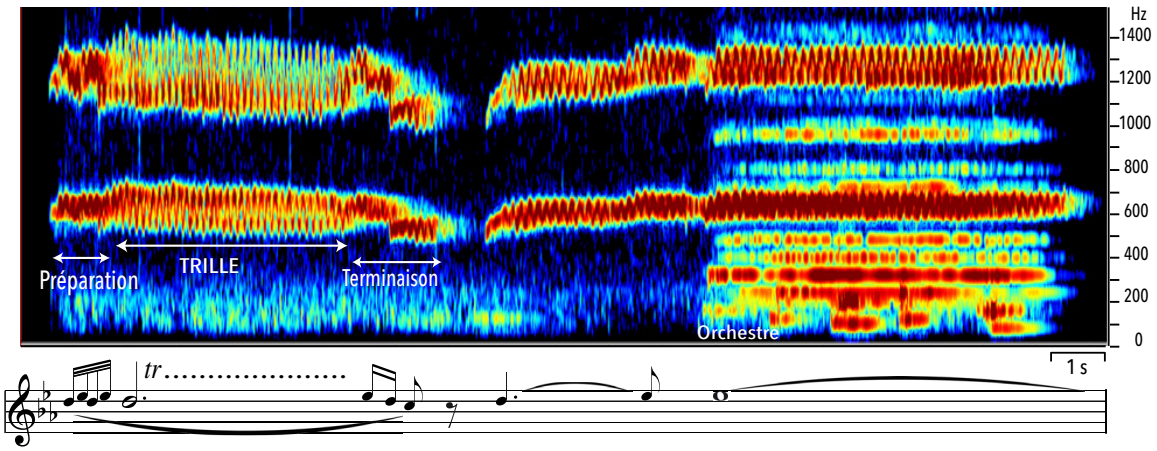


Figure 6.55 Analyse spectrographique d'un trille chanté avec préparation et terminaison.

4.5. À propos des ornements

Le vibrato de fréquence, aujourd'hui omniprésent dans l'interprétation de la musique classique et romantique, a supplanté la grande variété d'ornementations mélodiques qui se pratiquaient au temps de la monodie et que l'on peut apprécier dans les musiques traditionnelles, dans la musique dite populaire ou dans la variété.

Un grand nombre des ornementations instrumentales décrites dans les méthodes de musique de la Renaissance et de l'époque baroque sont, en quelque sorte, des adaptations instrumentales de la diversité des effets possibles à la voix.

Dans son traité sur la flûte traversière (1707), Hotteterre fait une grande place aux ornements du jeu. Le fait de boucher les trous de l'instrument avec les doigts offre une grande liberté dans la combinaison des trous ouverts ou fermés. À côté des « coulés, accents et ports-de voix » l'auteur décrit le flattement ou tremblement mineur et le battement de caractère plus énergique. Ce que nous nommons trille est appelé *cadence ou tremblement* – ce qui dit assez la délicatesse de l'ornement.

Quelques ornements ont été remis au goût du jour à la faveur de la redécouverte des musiques dites anciennes et traditionnelles. En témoigne le *bisbigliando* de la musique contemporaine, très répandu dans la pratique de la flûte *shakuhachi* japonaise. Produit par l'alternance rapide de la même note jouée avec des doigtés différents, il donne l'effet d'un tremblement de timbre.

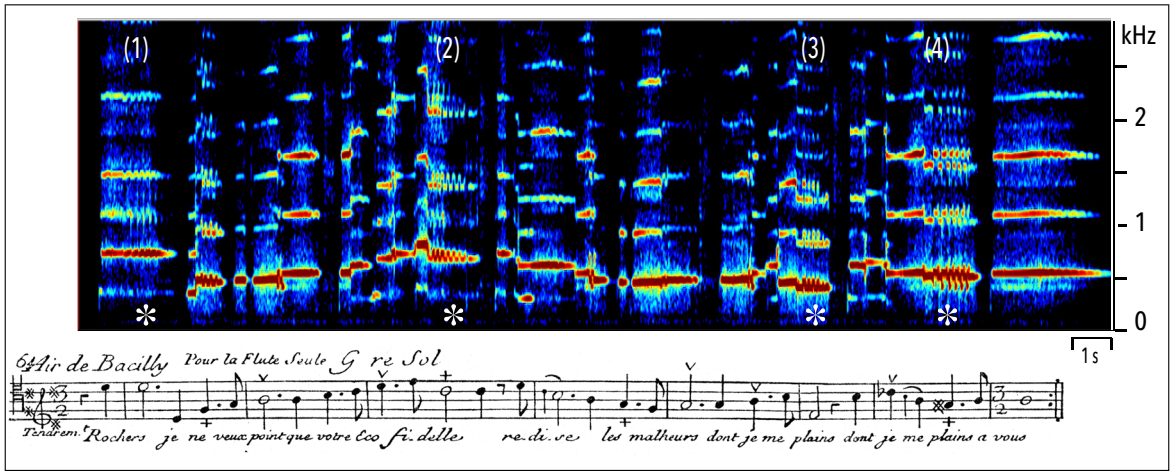


Figure 6.56 Analyse spectrographique d'une phrase musicale jouée à la flûte traversière à une clé. Notation musicale originale avec l'indication schématique des notes à orner (la croix pour le trille (ou cadence) ; le « v » pour un port-de-voix et battement.

Musique de Jacques Hotteterre, *Airs et brunettes*, s.d., p. 64 (collection privée).

La figure 6.56 présente l'analyse du jeu de la première phrase d'une mélodie de Bacilly (Son **6.68a**) transcrite par Hotteterre pour la flûte traversière à une clé. Quatre ornements sont repérés par un astérisque blanc. Le premier est un « flattement » de faible excursion (1/8 de ton). Les trois autres sont des « tremblements » dont l'effet change selon les notes affectées. La figure 6.57 montre l'analyse de deux d'entre eux plus en détail.

Flûte traversière à une clé (diapason la_3 = 415 Hz)

Son 6.68a (26″)

La première phrase de l'air « Rochers » de Bacilly (voir page 286)

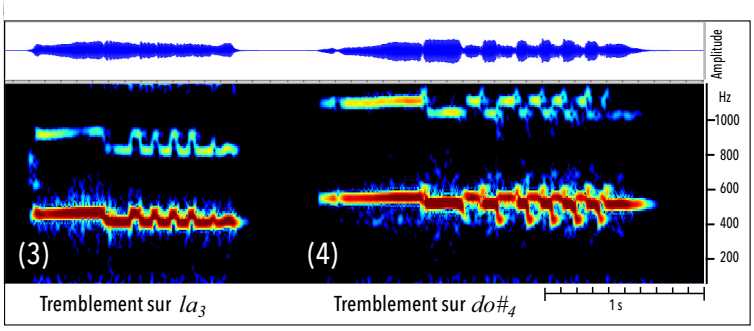


Figure 6.57 Exemples d'ornements à la flûte à une clé. (3) trille simple si_3-la_3. (4) trille $ré_4$-$do\#_3$ avec alternance de deux modes vibratoires différents : on entend – et on observe – des discontinuités mélodiques à chaque alternance de sons, accentuées par le contraste d'amplitude.

Le tremblement sur la_3 (3), alterne deux sons du même mode vibratoire dans une modulation continue. Le second (4) est caractéristique des effets sonores de ce type d'instrument (Son **6.68b**). Le $do\#_3$ et le $ré_4$ relevant de deux modes vibratoires différents, leur alternance rapide ne peut se faire sans une rupture de fréquence et des inflexions mélodiques qui changent selon le rang harmonique ; les discontinuités sont accentuées par les changements d'amplitude visibles sur la courbe supérieure.

Son 6.68b (7″)

Cadence (ou tremblement) sur la_3 et sur $do\#_4$

On voit que les modulations sonores produites par les techniques d'ornementation développées dans le jeu des instruments réels sont toujours complexes et infiniment variables. Les ornements possibles sur une flûte de facture baroque ont disparu avec les modifications de la perce et de la cléterie apportés par Boehm, au profit d'autres effets introduits par les modes de jeu contemporains (voir chapitre 7, page 337).

5. Bilan

À travers l'écoute et l'analyse des sons réels, la perception de la hauteur sonore, classiquement rapportée à la seule fréquence fondamentale des sons musicaux, apparaît d'une grande richesse et d'une grande diversité, tout en s'inscrivant dans les processus physiologiques présentés au chapitre 3.

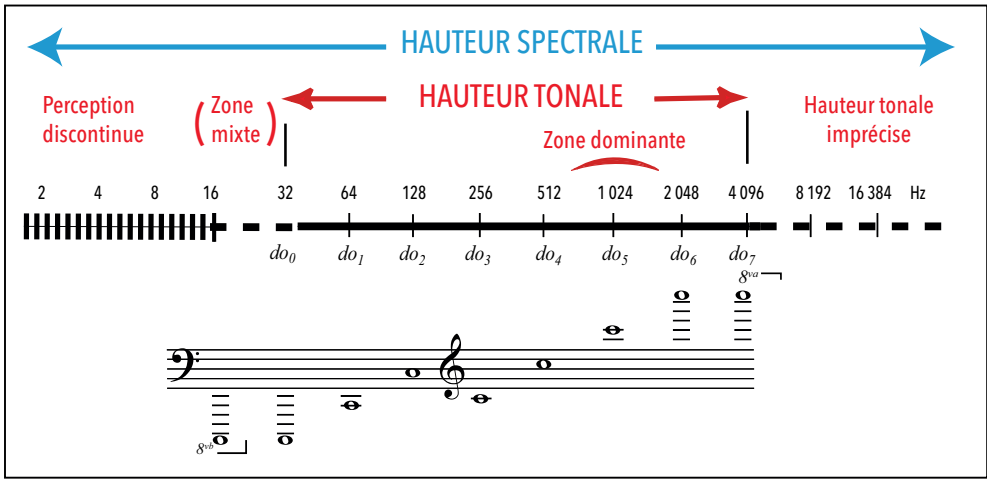


Figure 6.58 Tableau synthétique des capacités d'appréhension de la hauteur selon les fréquences. La zone d'écoute dominante, qui s'étend environ entre 800 et 1 800 Hz est distincte de la zone de la plus grande sensibilité de l'oreille (3000 Hz), non indiquée sur la figure.

5.1. La hauteur spectrale et la hauteur tonale

Les sons instrumentaux et vocaux ayant tous un spectre complexe, sollicitent simultanément les deux modalités de sensation de hauteur, la hauteur spectrale et la hauteur tonale.

Hauteur spectrale

La hauteur spectrale, due principalement au codage tonotopique, fournit une information globale sur la position des maxima énergétiques du spectre. Cette capacité primordiale de l'analyse auditive est active sur toute l'étendue du spectre des fréquences. De la succession de deux bruits dont les maxima de spectre diffèrent surgit une mélodie spectrale plus ou moins bien définie : mélodies de tambours, mélodies de ruisseaux (Son **6.2**). Dans le cas où il s'agit de sons périodiques pourvus d'une structure formantique, la mélodie suscitée par les variations spectrales est indépendante de celle qui naît des relations entre les hauteurs tonales (ou fréquences fondamentales) et peut donc entrer en conflit perceptif avec elle. Nous en avons donné de nombreux exemples (voir § 2.7). Toutefois, même lorsqu'elle s'impose perceptivement au point d'occulter la hauteur tonale, comme dans

l'exemple du contrebasson (Son **6.16**), la mélodie spectrale est imprécise : il est très difficile de la noter musicalement. Dans ce mode d'écoute, plus que la position précise des hauteurs, c'est la direction des variations spectrales qui attire notre attention, leur enveloppe temporelle ou contour mélodique.

Contour mélodique

La sensibilité au contour mélodique procurée par la hauteur spectrale est très commune et se développe vraisemblablement lors de l'apprentissage de la parole. On sait que les bébés sont sensibles au contour mélodique de la parole humaine qui est l'intonation, bien avant qu'ils soient capables d'en saisir le sens, et souvent les jeunes enfants chantent une comptine sans reproduire précisément les intervalles : seul le contour mélodique en est conservé.

Zone d'écoute dominante

Il est troublant également de remarquer que la zone de fréquences qui s'étend globalement de 700 à 1800 Hz, que nous avons appelée zone d'écoute dominante (figure 6.58) et qui correspond à peu de choses près à la zone d'existence des partiels dominants définie par Ritsma pour la sensation de hauteur (voir chapitre 3, figure 3.30), est aussi celle dans laquelle se produisent les variations du deuxième formant vocalique. Elle joue un rôle majeur dans l'appréciation de la hauteur des sons apériodiques (tubes, cloches).

Le fait que la hauteur spectrale ne donne pas d'information précise sur la grandeur des intervalles ne doit pas nous détourner de l'idée que cette compétence peut être cultivée à un point dont nous n'avons peut-être pas conscience, en particulier dans des traditions musicales qui paraissent, selon nos critères, dépourvues d'une échelle musicale repérable. La hauteur spectrale interviendrait aussi dans la discrimination mélodique des langues à tons et bien sûr dans l'écoute des musiques de rythmes.

Hauteur tonale

La hauteur tonale, principalement due au codage temporel, est celle qui permet d'apprécier de façon précise la hauteur de composantes périodiques stables, que les sons possèdent ou non de l'énergie à la fréquence fondamentale. Elle est aussi à invoquer pour la saisie de composantes isolées de certains sons apériodiques tels ceux des cloches. La hauteur tonale fonctionne dans un intervalle de fréquence limité, grossièrement entre 30 et 4000 Hz, zone qui correspond à la tessiture des fondamentaux musicaux, ce qui n'est pas surprenant si l'on songe que les instruments se sont adaptés au fil des siècles pour exploiter au mieux les capacités de l'oreille humaine. Au-delà de 4000 Hz, nous percevons toujours une hauteur, mais sans pouvoir précisément définir les intervalles de variation. Du côté des basses fréquences, la sensibilité aux variations temporelles quitte le domaine des hauteurs pour celui des rythmes. Nous sommes extrêmement sensibles aux variations lentes de la fréquence des sons, aux ornements de toute sorte qui donnent la vie aux sons produits par les gestes humains. Dans la zone intermédiaire 10-30 Hz, les sensations varient selon les auditeurs, mêlant les aspects discontinus et continus de la vibration (voir aussi chapitre 3, § 3.2).

Développée à l'écoute des sons d'instruments de musique entretenus, riches en harmoniques, la hauteur tonale permet d'atteindre une très grande précision dans la discrimination des hauteurs, comme le révèlent les seuils différentiels (voir chapitre 3, § 2.5) : jusqu'au 1/1000 de la fréquence pour un son de 1000 Hz ! Contrairement à la hauteur spectrale, la hauteur tonale est une capacité qu'il faut acquérir et entraîner. Spécifique de la musique, l'évaluation des intervalles fait appel à des traitements très précis et suppose des apprentissages complexes pour l'écoute de la polyphonie et de l'harmonie. Le cas particulier du vibrato

(voir l'exemple du Son **6.54**) et son expansion en musique vocale montrent l'importance de la part culturelle et éducative dans l'appréciation de la hauteur des sons.

Apprendre à repérer la hauteur tonale liée à la périodicité du signal, c'est apprendre à repérer ce qui est commun à deux sons produits sur des instruments différents, quelles que soient les variations de hauteur spectrale. De ce point de vue, hauteur spectrale et hauteur tonale seraient donc antagonistes et souvent conflictuelles (sons **6.18** à **6.21**).

La précision qu'offre la hauteur tonale pour apprécier la hauteur des sons périodiques porte sur l'évaluation des rapports entre les sons et non sur la mesure d'une fréquence. Un son isolé n'est pas de la musique. C'est de la comparaison entre sons successifs ou sons simultanés que naît l'intervalle. Autrement dit, la hauteur tonale est l'aptitude que développent prioritairement les musiciens, et en particulier ceux qui ont l'oreille dite relative[31].

5.2. La perception de la hauteur, la facture instrumentale et les modes vibratoires

Dans ce chapitre sur la perception de hauteur nous avons évoqué à plusieurs reprises les modes vibratoires des instruments de musique dont l'ajustement est sous le contrôle du facteur d'instrument. Voici une petite synthèse.

Sons périodiques

- Pour les sons périodiques, on utilise un seul mode, le plus souvent le premier sur les instruments à cordes (excepté les modes de jeu en sons « harmoniques »), les trois ou quatre premiers modes du tuyau pour les instruments à trous latéraux, selon le registre, et tous les modes jouables aux cuivres. De toute façon, comme le son est entretenu, la hauteur de la note jouée est parfaitement définie. La justesse des modes entre en ligne de compte pour l'extension de la tessiture (afin de pouvoir octavier, quintoyer) et elle a une incidence importante sur la richesse spectrale du son joué.

Sons apériodiques

- Pour répondre au critère d'une hauteur bien définie, accordable, les sons apériodiques des instruments non entretenus requièrent des réglages spécifiques. Le premier groupe est celui des instruments à cordes libres. Les fréquences des modes propres doivent être très voisines de celles de la série harmonique du mode 1 pour que l'oreille s'en accommode et leur attribue une fondamentale. À l'opposé se situe le groupe des instruments pour lesquels un seul mode est favorisé (verges libres). Les autres modes sont neutralisés (point de frappe, point d'appui) ou accordés sur une partie aliquote du mode 1. Le troisième groupe (cloches, timbales) rassemble les instruments pour lesquels un groupe de modes propres consécutifs (2 ou 3) sont de bons candidats au modèle harmonique. L'accordage et le jeu (point de frappe) contribuent alors à favoriser la note désirée.
- Lorsqu'il est possible de sélectionner et de stabiliser plusieurs modes vibratoires – cas des sons multiphoniques – le son produit donne lieu à la perception de plusieurs hauteurs distinctes formant un accord complexe de timbres et de hauteurs (voir le chapitre 8, § 2.4.2).

Son 6.69 (10")

Musique jouée au marimba, dans le grave de la tessiture

Dans tous les cas, la perception de la hauteur des sons apériodiques dépend étroitement de la composition spectrale et donc varie fortement avec la tessiture. La note au coup des cloches n'est bien perçue que dans un intervalle d'une octave et demie. La note des claviers (marimba, vibraphone et xylophone) est difficile à percevoir dans le grave en raison des partiels supérieurs, très présents (Son **6.69**), mais la tessiture peut s'étendre sans problème dans l'aigu.

31. Voir le chapitre 8, page 435, pour quelques réflexions sur l'oreille dite absolue.

5.3. Perception de la hauteur et mesure acoustique

On voit qu'il est capital de connaître très précisément la structure acoustique des sons que l'on étudie sous l'angle de la hauteur, surtout si l'on souhaite y faire correspondre les résultats d'une mesure acoustique. Avant de procéder à une analyse de la hauteur fondamentale, deux questions préalables sont requises :

- s'agit-il d'un son périodique ? À titre de rappel, voici la même note, jouée sur un vibraphone, d'abord percutée, puis entretenue à l'archet (Son **6.70**) : à l'écoute, ces deux sons appartiennent bien à des mondes sonores radicalement différents (voir aussi l'exemple du verre frappé et frotté, chapitre 1, Sons **1.1** et **1.2**).
- dans quelle tessiture se trouve-t-il ? L'enveloppe spectrale pourra donner une indication sur la situation du maximum énergétique, c'est-à-dire l'interférence possible de la hauteur spectrale avec la hauteur fondamentale. S'il s'agit de sons apériodiques, l'analyse à l'oreille et le travail de transformation par synthèse sont nécessaires pour interpréter la perception de hauteur, sans oublier le contexte et les particularités individuelles qui ont une influence importante sur la perception.

Son 6.70 (6")

Lame de vibraphone percutée puis frottée par un archet

5.4. La hauteur comme qualité de sons connus et catégorisés

Seul l'apprentissage permet finalement de comprendre que nous puissions attribuer la même hauteur à des sons aussi différents que ceux du piano, de la cloche tube, du marimba ou de la flûte : lorsque cet apprentissage fait défaut nous sommes démunis (Son **6.6**). Ceci nous ramène à la caractéristique fondamentale de la perception humaine dont le premier réflexe est d'identifier la source sonore. Loin d'interférer, voire de contrecarrer la perception de la hauteur, nous pensons que « la reconnaissance du timbre identitaire est une étape primordiale permettant d'adapter efficacement les stratégies d'écoute en vue de l'interprétation de la hauteur musicale ».

Pour chaque catégorie de sons, nous mémorisons des critères spécifiques qui s'affinent au cours de la pratique musicale. Un exemple frappant est celui des cloches dont on « apprend » réellement à écouter la hauteur[32]. Un autre est celui des mélodies spectrales du chant harmonique, que plusieurs musiciens entraînés à la saisie des « notes » ne parviennent à entendre qu'après une adaptation de leur écoute afin d'ignorer la hauteur fondamentale et les voyelles.

C'est aussi la connaissance de la source et de ses caractéristiques qui permettent de comprendre que la moindre fluctuation de hauteur soit inacceptable pour un son de piano – car nous « savons » que cet instrument a une hauteur fixe – et qu'à l'opposé nous puissions entendre « une note » alors que le chanteur fait un vibrato balayant un intervalle de plus d'un demi-ton !

Le lecteur pourra croiser avec profit les résultats énoncés dans ce chapitre avec ceux que nous développons dans le cadre de l'étude du timbre (chapitre 7), ainsi que dans le chapitre 8 dans lequel nous abordons l'étude des intervalles musicaux et celle des phénomènes se produisant lors de l'émission de sons simultanés.

32. L'auteur se souvient avoir réagi vivement au désagrément procuré par l'écoute d'une pièce de musique classique de tonalité majeure interprétée sur un carillon réputé, alors qu'aujourd'hui cette audition lui procure un plaisir certain !

Réponses à la question posée page 240 – consigne d'écoute du Son 6.16

- Réponse 1 : dessin d'une ligne brisée irrégulière.
- Réponse 2 : dessin d'un escalier ascendant.

Comment expliquer une telle divergence de perception à l'écoute d'un instrument de l'orchestre ?

- Les auditeurs qui dessinent un escalier ascendant perçoivent la hauteur tonale, c'est-à-dire les notes de la gamme ascendante effectivement jouée, qui se situe dans la zone du do_0.
- Les auditeurs qui dessinent une ligne brisée accidentée perçoivent la hauteur spectrale d'une mélodie formantique qui se situe dans l'octave do_3-do_4. Cette mélodie résiste à une notation précise sur portée et varie d'un auditeur à l'autre.

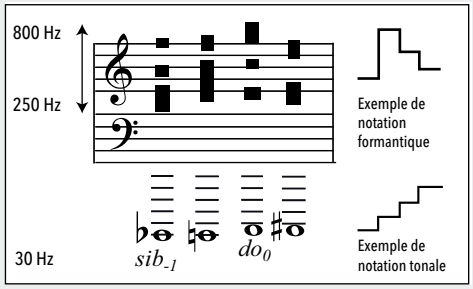


Figure 6.59 Notation sur portée de la hauteur tonale et figuration des zones formantiques. En vis-à-vis, exemples des deux types de dessins mélodiques recueillis lors de l'expérience d'écoute.

6. Les sons du chapitre 6

6.1. Trois expériences introductives

Son 6.1 – Expérience n° 1. Quatre sons brefs, extraits de l'environnement sonore : écoute causale, orientée sur la reconnaissance des sources. [M. C.]

Son 6.2 – Chacun des quatre sons précédents a subi deux transpositions. Regroupement des sons par source d'origine : écoute qualitative orientée sur les modifications spectrales produites par la transposition. Remarque : la transposition en fréquence s'accompagne d'une modification de la durée qui n'a pas d'incidence sur l'expérience proposée. [M. C.]

Son 6.3 – Expérience n° 2. Estimation d'un intervalle de hauteur entre trois couples de sons. La transposition appliquée à chaque type de son est strictement identique, alors que l'intervalle perçu entre deux sons successifs varie notablement. Seul le son harmonique de flûte à bec permet une appréciation précise de la fréquence. [M. C.]

Son 6.4 – Expérience n° 3. Quatre sons émis par des sources différentes donnant la même perception de hauteur. [M. C.]

6.2. Sons périodiques

6.2.1. Hauteur perçue et connaissance de la tessiture d'une source

Son 6.5 – Test d'appariement d'un son vocal avec un son sifflé. Présentation de deux paires. Y a-t-il unisson ? [M. C.]

Son 6.6 – Test d'appariement voix et sifflet : présentation de trois paires. Seul le sifflet du dernier exemple est à l'unisson de la voix. [M. C.]

6.2.2. Tessiture et hauteur tonale

Son 6.7 – Tuba : quatre notes tenues, détachées, jouées lentement. Essayez de noter les intervalles de cette succession. [M. C.]

Son 6.8 – Tuba : la séquence entière de 12 notes (gamme chromatique ascendante)[33]. Identifiez la note à partir de laquelle vous percevez une succession chromatique. Bref silence ; on entend ensuite les huit premiers sons transposés à l'octave supérieure (fréquences doublées). Les intervalles sont bien perçus par tous au-delà de 40 Hz. [M. C.]

Son 6.9 – Trois séquences de voix parlée : extrait de poésie. a) voix normale ; b) voix ralentie et transposée à l'octave inférieure : la hauteur fondamentale de la voix disparaît au profit d'une suite d'impulsions ; c) de nouveau la voix normale. Voix d'homme grave (fréquence fondamentale entre 70 et 90 Hz). Début du poème *Le Cimetière marin* de P. Valéry, dit par Jean Vilar ; date inconnue. [Archives LAM]

6.2.3. Expériences de filtrage

Son 6.10 – Phrase mélodique jouée au basson : trois filtrages puis la séquence originale. Son 6.10a, filtre passe-haut (1000 Hz) ; Son 6.10b, filtre passe-bande (450-1850 Hz) ; Son 6.10c, filtre passe-bas (1000 Hz) ; Son 6.10d, son original. Allegro du *Concerto en* sib *majeur* de Mozart ; Maurice Allard, 1965. [Archives LAM]

Son 6.11 – Basson. La première note (do_3 = 270 Hz) de chacune des quatre séquences de l'exemple précédent. Notez que la sonorité se modifie, mais que la hauteur de la note n'est pas affectée. L'analyse spectrale des quatre sons est visible figure 6.9 ; a) est privé des trois premiers harmoniques ; b) contient les harmoniques 2 à 6 ; c) contient les harmoniques 2 à 4 ; d) est le son original. [M. C.]

Son 6.12 – Qualité sonore et technique d'enregistrement. Son 6.12a : sonnerie de trompes de chasse ; disque 78 tours lu par un appareil mécanique (bande passante 500-4000 Hz) ; Son 6.12b : même disque lu par un appareil électromécanique ; Son 6.12c : sonnerie de trompes de chasse enregistrée sur bande magnétique avec un Nagra. Paris, place de la Madeleine, 1970. [Archives LAM]

6.2.4. Hauteur tonale et spectre (sonorité)

Son 6.13 – Sons synthétisés au Cantor[34], instrument permettant de modifier le spectre harmonique sans altérer la fréquence fondamentale. a) la note fa_3 jouée trois fois : spectre pauvre avec fondamental dominant ; spectre riche avec fondamental faible ; spectre riche avec les cinq premiers harmoniques dominants ; b) même expérience effectuée à l'octave supérieure, fa_4. [M. C.]

Son 6.14 – Incidence du point de pincement sur la perception de hauteur du son d'une corde grave (60,9 Hz). Son 6.14a : une corde de Nylon est pincée au milieu de sa longueur (2 fois) ; Son 6.14b : la même corde est pincée près du sillet (2 fois). Ces deux sons ont-ils la même hauteur ? Certains auditeurs perçoivent deux notes différentes alors que la fréquence fondamentale de la corde est strictement la même. Rappel : la corde pincée au milieu ne produit que les composantes de

33. Fréquences des 4 premières notes : $fa\#_{-1}$ (22,78 Hz) ; sol_{-1} (24,80 Hz) ; $sol\#_{-1}$ (26,44 Hz) ; la_{-1} (27,75 Hz).
34. Orgue électronique construit au LAM.

rang impair. La perception de cet exemple est très dépendante de la qualité sonore d'écoute. [M. C.]

6.2.5. Diapason et sonorité

Son 6.15 – Effet du diapason sur les qualités du timbre d'un orgue. Son 6.15a : enchaînement de quatre séquences. 1) la première mesure (4'') de l'enregistrement effectué en été (diapason haut) ; 2) la première mesure de l'enregistrement effectué en hiver (diapason bas) ; 3) et 4) : les deux enregistrements précédents mis au même diapason (durée de la phrase un peu plus longue). Son 6.15b : la phrase complète enregistrée en été. Son 6.15c : la phrase complète enregistrée en hiver. Remarque : les conditions d'enregistrement (matériel d'enregistrement, position par rapport à l'instrument, registration, etc.) sont strictement identiques. Orgue de Saint-Maximin-en-Provence, avant restauration ; mission du ministère de la Culture, 1986, CD 2. Début du Fond d'orgue de Nivers, M. C. [Archives LAM]

6.2.6. Hauteur tonale et hauteur spectrale : cas des sons graves

Son 6.16 – Contrebasson. Les cinq premières notes d'une mélodie : notez ce que vous entendez. [M. C.]

Son 6.17 – Contrebasson. La séquence complète (gamme chromatique ascendante) et sa transposition à l'octave supérieure (fréquences doublées). [M. C.]

6.2.7. Hauteur tonale et hauteur spectrale : paradoxes

Son 6.18 – Séquence de sons synthétisés à l'ordinateur produisant une gamme descendante ininterrompue. [© R. Shepard, 1964]

Son 6.19 – Orgue à tuyau. Gamme chromatique descendante jouée sur le registre de Cymbale. Les « reprises » produisent un effet similaire à l'exemple du son précédent : une descente continue qui semble se maintenir dans la même tessiture. Orgue de Graffenstadten, 1966 [M. C.].

Son 6.20 – Synthèse à l'ordinateur d'un glissando descendant sans fin. [J.-C. Risset]

Son 6.21 – Synthèse à l'ordinateur d'un son paradoxal faisant entendre les croisements d'une hauteur tonale descendante avec une hauteur spectrale ascendante. [J.-C. Risset]

6.2.8. Hauteur tonale et hauteur spectrale : effets vocaux

Son 6.22 – 1) Glissando vocal ascendant-descendant sur la voyelle « a » : hauteur fondamentale variable et formants vocaliques constants. 2) Succession de cinq voyelles « a, e, i, o, u » enchaînées recto-tono : hauteur tonale constante et formants vocaliques variables. Voix masculine ; B. Roubeau, 1996. [D. Mercier]

Son 6.23 – Exemple de changements vocaliques : « ou, i, ou, i » sur une note tenue ($ré_3$). La hauteur tonale est constante et on entend un glissando spectral ascendant et descendant dû à la variation du second formant vocalique. Voix féminine. [M. C.]

Son 6.24 – En articulant « i, ou » avec une intonation ascendante, on peut croiser la hauteur tonale et la hauteur spectrale. L'effet s'inverse sur « ou, i » avec une intonation montante. Voix féminine. [M. C.]

Son 6.25 – Chant harmonique. Exemple pédagogique d'un glissando ascendant-descendant. À la suite, même exemple ralenti et transposé à l'octave inférieure (fréquences divisées par deux et durée doublée). Trân Quang Hai, 1987. [Archives LAM]

Son 6.26 – Arc musical. Mélodie harmonique avec deux fondamentaux à un ton d'intervalle. Musicien pygmée ngbaka enregistré par S. Arom (1967). [Archives LAM]

Son 6.27 – Guimbarde et chant harmonique de style sygyt. Successivement : jeu de la guimbarde (fondamental à 88,5 Hz) ; 13'' guimbarde et voix (à l'octave supérieure de la guimbarde) ; 30'' deuxième séquence voix et guimbarde ; 45'' guimbarde seule. *Tuva, Voices from the Center of Asia* ; n° 16 ; Smithsonian Folkways Records, 1990. [SF 40017]

Son 6.28 – Chant harmonique de style kargiraa. Note fondamentale très grave, do_0 = 65,5 Hz produite par *period-doublin*g. Dans ce style de chant, l'écoute peut se porter alternativement sur la mélodie ou sur les voyelles. *Ibid,* n° 1. [*Ibid*.]

Son 6.29 – Chant harmonique xhosa à deux fondamentaux. Note fondamentale grave en *period-doubling*. Style *umngqokolo ngomqangi*[35]. Mme Mbizweni, Dashe, village Ngqoko (Afrique du Sud), 1985. [D. Dargie]

Son 6.30 – Chant tibétain. Exemple d'unisson spectral et tonal. Au terme d'une montée très progressive, la fréquence fondamentale des moines du chœur fusionne avec l'harmonique 5 du soliste. Montage expérimental de courtes séquences prélevées dans la séquence complète. *The Heart of Dharma*, plage 2, *Gaden Lhagyama ;* 1996. [Ellipsis arts]

6.3. Sons apériodiques

6.3.1. Sons quasi périodiques

Son 6.31 – Deux sons de piano dont on n'entend que la partie aiguë du spectre, les composantes graves ayant été filtrées. Le deuxième son (numérisation d'un son réel de piano) est généralement perçu plus aigu, d'environ un demi-ton. Taro Mori, 2000, thèse. [Archives LAM]

Son 6.32 – Les deux sons de l'exemple précédent, avant filtrage. Le premier son est périodique : toutes les composantes sont des multiples entiers de la première composante (la plus grave). Le deuxième son, inharmonique, est un son réel de piano. Les deux sons donnent la même *note*, mais ont une sonorité différente. Taro. Mori. [*Ibid*.]

6.3.2. Hauteur d'une composante dominante

Son 6.33 – Diapason à fourche. L'instrument est frappé puis posé sur une table d'harmonie. Le son du premier mode vibratoire – le la_3 proprement dit – est accompagné d'un son beaucoup plus aigu, perceptible à la frappe, qui est le second mode vibratoire du système. [M. C.]

Son 6.34 – La note do_2 jouée sur un marimba basse[36]. Base McGill. [M. C.]

Son 6.35 – Comparaison de la note do_3 jouée sur un marimba puis sur un vibraphone. Base McGill. [M. C.]

6.3.3. Le modèle harmonique

Son 6.36 – Cloche tube. Son du tube accordé sur do_3. [Base McGill]

Son 6.37 – Cloche tube. Le son précédent suivi d'un son sinusoïdal ayant pour fréquence le do_3. [M. C.]

Son 6.38 – Analyse auditive des composantes isolées de la cloche tube do_3. On entend successivement : le son global de la cloche tube ; (4''5) la suite ascendante des partiels n° 2 à n° 9 ; (31'') le son global de la cloche tube. [M. C.]

35. *ngomqangi* est le nom de l'arc musical xhosa.
36. Si, à l'écoute, le son de cet exemple vous semble plus aigu que les deux suivants, c'est que votre système de reproduction filtre le fondamental du marimba.

Son 6.39 – La cloche tube précédente avec un filtrage passe-bas (coupure à 1070 Hz) ne conservant que les partiels 2, 3 et 4 qui sont responsables de la hauteur perçue, suivie du son sinusoïdal de 263 Hz (do_3). Le timbre est modifié, mais la « note » de la cloche est bien celle du son sinusoïdal. Il est intéressant de faire la comparaison avec le son original 6.36. [M. C.].

Son 6.40 – Sonnerie d'une cloche de volée. Repérez les différentes notes perceptibles. Cloche de Hochfelden, 1981. [M. C.]

Son 6.41 – Synthèse additive de la cloche précédente permettant de modifier les fréquences des partiels. Son 6.41a : dans cet exemple, les deux partiels évoquant la quarte ont été accordés sur des composantes harmoniques de la première note au coup ; la seconde note au coup a disparu ; Son 6.41b : les deux partiels sont à leur fréquence réelle : la note de quarte réapparaît. [M. C.]

6.3.4. La zone d'écoute dominante

Son 6.42 – Expérience de transposition en fréquence d'une cloche. Son 6.42a : la cloche de l'exemple 6.40 est jouée normalement ; Son 6.42b : la cloche est transposée à l'octave supérieure. Toutes les fréquences sont multipliées par deux, de même que la cadence des coups. La seconde note au coup (la quarte) a disparu au profit de la tierce mineure qui devient prédominante. En décalant les partiels par rapport à la zone d'écoute dominante, cette transposition transforme fondamentalement la perception du son de la cloche. [M. C.]

6.3.5. La note d'une timbale (à écouter avec de bonnes enceintes)

Son 6.43 – Une note de timbale d'orchestre. Dans quelle octave se situe la note de cette timbale ? fa_1 ou fa_0? [M. C.]

6.3.6. Sons inharmoniques successifs : le plus court trajet mélodique

Son 6.44 – Deux sons d'un jeu de cloche tube. L'intervalle perçu (le plus souvent une tierce majeure) correspond rarement à celui qui est joué : ici, c'est l'octave do_3-do_4. Sons de la base McGill. [M. C.]

Son 6.45 – Son inharmonique synthétisé à l'ordinateur. L'écoute successive du son et de sa transposition (toutes les fréquences doublées) ne produit pas une translation à l'octave supérieure ! L'intervalle perçu est généralement un demi-ton descendant. [J.-C. Risset]

Son 6.46 – Cloche tube. La suite des 13 tubes donnant les notes d'une gamme chromatique de do_3 à do_4. Après le silence (vers 41 secondes), le treizième tube (do_4) est répété puis suivi du premier tube (do_3). Dans la rémanence auditive de la gamme précédente, ces deux tubes sont maintenant perçus à l'octave l'un de l'autre. Sons de la base McGill. [M. C.]

Son 6.47 – Cloche à deux « notes » (deux points de frappe). Son 6.47a : sonnerie de la cloche à la volée. On entend initialement la cloche frappée par le battant intérieur ; après une douzaine de secondes, le son de la tierce mineure, dû au marteau roulant extérieur, s'intercale entre les coups du battant ; Son 6.47b : écoute de la zone 500-1200 Hz : la cloche donne toujours à entendre les deux notes, ce qui confirme l'interprétation proposée. Cloche de la commune de Ry, en Normandie ; mission des Affaires culturelles, 1971. [Archives LAM, M. C.]

6.3.7. Sons inharmoniques fortement amortis

Son 6.48 – Percussion sur deux tables d'harmonie d'un instrument en cours de construction (luth). De la comparaison des chocs sur chaque table naît un intervalle musical : ascendant ? descendant ? ou plus complexe ?[37] [M. C.]

37. Pour cet exemple aussi, les réponses sont dépendantes du système d'écoute que vous utilisez (enceintes de salon, casque, sortie son de l'ordinateur).

6.3.8. Exemple de filtrage d'une mélodie jouée sur des tom-tom : exemple à prendre en dictée

Son 6.49 – Écoute d'une formule mélodique jouée sur trois tom-tom. L'exemple propose sept présentations de la séquence musicale entendue au travers d'une suite de filtrages passe-bande, de l'aigu au grave. Dans chaque bande, la mélodie est reproduite deux fois, puis après environ 6 secondes, on passe à la bande suivante. À chaque présentation, notez sommairement la mélodie perçue, ou tout au moins le début. La mélodie de 12 sons est jouée sur 3 « notes ». La dernière bande correspond au son global, c'est-à-dire au son non filtré. 1970. [E. Leipp]

6.4. Instabilités, vibrato, trille

Son 6.50 – Trois sons (la_3) joués *droits*, c'est-à-dire avec la plus grande stabilité possible. Successivement : flûte à bec, violon, voix. [M. C.]

6.4.1. Vibratos divers

Son 6.51 – Chanteur avec vibrato d'amplitude et différents vibratos de fréquence ; tradition syrio-libanaise. Disque *Les Voix du monde*, collection musée de l'Homme/CNRS ; Harmonia Mundi, CD III n° 19 ; 1996. [Document Laurent Aubert]

Son 6.52 – Trois sons vibrés sur la note do_3. Son 6.52a : un vibrato joué à la flûte traversière ; Son 6.52b : un vibrato joué au violon ; Son 6.52c : un vibrato de voix chantée (basse). Base McGill (flûte, violon) ; base RWC (voix extraite du fichier 493BSE3M). [M. C.]

Son 6.53 – Synthèse d'un son vocal vibré sur la voyelle « o ». La rapidité du vibrato est constante, on varie l'intervalle ($\Delta f/fm$). Présentation alternée du son vibré (V) puis du son fixe (F) à la fréquence moyenne *fm*. Deux alternances « VFVF ». Trois largeurs de vibrato, successivement : a) $\Delta f/fm$ = 3 % (1/4 de ton) ; b) $\Delta f/fm$ = 6 % (1/2 ton) ; c) $\Delta f/fm$ = 12 % (1 ton). Dans l'exemple c), la hauteur moyenne perçue « bouge » avec le vibrato. Pour certains auditeurs, elle monte, pour d'autres, elle descend. [M. C.]

Son 6.54 – Vibrato et tolérance de justesse. Arpège ascendant (do_3, sol_3, do_4) dont les sons 2 et 3 sont vibrés (deux fois). Ensuite, même arpège sans vibrato, c'est-à-dire que les sons émis sont les fréquences centrales des sons précédemment vibrés ! La fausseté est soudain insupportable. [M. C.]

6.4.2. Vibrato vocal des chanteurs

Son 6.55 – Exemple d'un grand intervalle chanté avec vibrato. L'exemple commence sur $ré_3$. Quelle est la note aiguë ? Écoutez plusieurs fois pour noter par écrit l'intervalle musical[38]. [Références du Son 6.57]

Son 6.56 – Vibrato vocal sur une gamme rapide (une « arche » de vibrato par note). La gamme diatonique do_3-sol_4 chantée par une soprano (sens ascendant et descendant), suivie de la même gamme jouée au piano. [LAM]

Son 6.57 – Transition liée et vibrée, sur un grand intervalle musical. L'extrait centré sur le grand intervalle de la transition analysée figure 6.49 se situe entre la 6e et la 9e seconde. Rossini, *La Donna del lago*, rondo final d'Elena, M. Horne. [Archives LAM]

Son 6.58 – Arpèges en notes piquées vibrées. Mozart, *La Flûte enchantée*, extrait de l'air de la Reine de la nuit, W Lipp. [Archives LAM]

Son 6.59 – Synthèse de notes courtes vibrées. On entend successivement les quatre portions différentes de vibrato, c'est-à-dire les formes a, b, c, d de la figure 6.51. Synthèse par FOF. [C. D'Alessandro, M. C.]

38. Les mesures (Praat) de la fréquence moyenne des sons vibrés sont : 297 Hz pour le grave ; 852 Hz pour l'aigu, soit un grand triton ou une petite quinte selon les auditeurs.

Son 6.60 – Exemples d'un test d'ajustement de la hauteur perçue. Son 6.60a : succession rapide de cinq paires de notes très courtes. Le 1er son de chaque paire, X, est vibré (forme a), et le second, Y, est un son fixe de fréquence variable que le sujet ajuste pendant le test. Les fréquences des cinq sons fixes sont successivement : Y = 430, 438, 455, 446, 450 Hz. La fréquence moyenne de la forme a est 450 Hz (voir figure 6.51) ; Son 6.60b : son vibré (forme b) de durée croissante ; la hauteur perçue ne change pas. Les quatre durées sont successivement : 80 ms, 120 ms, 160 ms et 330 ms. [C. D'Alessandro, M. C.]

6.4.3. Le cas étrange du trille vocal

Son 6.61 – Exemple d'un trille vocal suivi d'un trille à la flûte traversière (octave supérieure). Donizetti, *Lucia di Lammermoor* ; Mado Robin, Decca, LXT 2898. [Archives LAM]

Son 6.62 – Synthèse d'une voix chantant les trois sons liés, *la*$_3$, *si*$_3$, *do*#$_4$, avec un vibrato de largeur croissante. Synthèse par FOF. [M. C.]

Son 6.63 – Synthèse d'une voix chantant un trille avec la préparation (*do*$_3$, *ré*$_3$). Synthèse par FOF. [M. C.].

Son 6.64 – Écoute des extraits (1) et (2) de la figure 6.53, qui sont acoustiquement identiques mais perçus différemment : tantôt comme un vibrato (Son 6.62), tantôt comme un trille (Son 6.63). [M. C.]

Son 6.65 – Deux trilles vocaux isolés du contexte musical : quel intervalle fait le premier ? le second ? [M. C.]

Son 6.66 – La phrase musicale du premier trille (1/2 ton) ; la phrase musicale du second trille (1 ton). [M. C.]

Son 6.67 – La séquence complète d'un trille vocal en contexte, avec préparation et finale. J. Sutherland. Bellini, *La Sonnambula,* air « Come per mer sereno ». Disque *The Voice of the Century*, London Records,STERO OSA 13107, 1962. [M. C.]

Son 6.68 – Ornements spécifiques à la flûte traversière à une clé. Son 6.68a : jeu d'une mélodie ornée, « Rochers » de Bénigne de Bacilly ; Son 6.68b : détail de deux ornements extraits de la phrase musicale : le trille *si*$_3$-*la*$_3$ et le trille *ré*$_4$-*do*#$_3$. Flûte Martin Lot, musée du CNSM. Pierre Séchet, 1978. [Archives LAM]

Son 6.69 – Une petite mélodie jouée dans la première octave grave d'un marimba (*fa*$_1$-*fa*$_2$). Dans cette tessiture, les partiels des modes supérieurs deviennent dominants perceptivement, et perturbent l'écoute de la hauteur. Seule une reproduction sonore de qualité (enceintes graves) restitue la mélodie jouée. CNSM, 1997. [M. C.]

Son 6.70 – La lame d'un vibraphone (note *sib*$_2$) jouée successivement de deux façons différentes : percutée avec une baguette, entretenue avec un archet. Base McGill. [M. C.]

CHAPITRE 7

LA QUESTION DU TIMBRE

1. La musique et le matériau sonore

1.1. Le timbre, la musique et l'acoustique

La notion de timbre en musique apparaît au XVIIIe siècle dans la culture occidentale. À cette époque, la musique est polyphonique, écrite par un compositeur pour un groupe d'instruments quasi normalisés pouvant être combinés de diverses façons au sein d'un ensemble à géométrie variable : l'orchestre. Cette notion se constitue au moment même où le son, « matériau » de la musique, s'autonomise comme objet d'étude scientifique. Le timbre est alors défini comme la troisième caractéristique du son, conjointement avec l'intensité et la hauteur, et toujours sur la base d'exemples empruntés aux instruments de musique. Aujourd'hui, les diverses définitions du timbre et les études de psychoacoustique s'inscrivent peu ou prou dans la même tradition. Pourtant, le travail du timbre n'a cessé de jouer un rôle de plus en plus important dans la structure même des œuvres musicales, certains compositeurs ayant même acquis le statut de *créateurs de timbre,* en s'appropriant les moyens modernes de génération et de transformation des sons, au même titre que les instrumentistes. De leur côté, les acousticiens ont bénéficié des moyens offerts par la synthèse sonore numérique pour explorer les dimensions perceptives du timbre.

Toutefois une importante ambiguïté subsiste : qu'entend-on exactement par ce mot « timbre » ? Musiciens et acousticiens ont fait part de leurs réflexions dans de nombreux ouvrages[1]. La lecture de ces différents textes produit tout d'abord une impression de perplexité tant il est difficile de mettre en relation l'expérience riche et foisonnante des compositeurs avec celle des acousticiens qui travaillent généralement avec des sons isolés, hors de tout contexte musical. Indépendamment des différences de contexte, artistique ou scientifique, il semble bien que, sous le même terme, les deux communautés ne visent pas le même objet d'étude ! Contrairement aux autres caractéristiques du son – hauteur, intensité, durée –, le timbre n'est pas assimilable à une grandeur quantifiable. D'ailleurs, il n'existe pas de notation musicale du timbre : c'est un problème auquel sont confrontés les compositeurs actuels.

La déception des compositeurs qui ont tenté de créer des timbres par synthèse sonore n'a d'égal que le désarroi des acousticiens qui, au terme de soixante années de recherches en psychoacoustique, font état du faible avancement de leurs

1. Citons P. Schaeffer (1952, 1966), G. Ligeti (1975), R. Erickson (1975), J.-C. Risset (1994, 2004) pour la composante musique ; R. Plomp (1976), J. Grey (1977), W. Slawson (1985), J.-C. Risset & D. Wessel (1982), J. Beauchamp (2006) pour l'étude acoustique, ainsi que l'ouvrage édité à la suite d'un colloque pluridisciplinaire organisé par J.-B. Barrière, dans lequel on trouve des contributions d'auteurs divers. (Collectif – Barrière, *Le timbre : métaphore pour la composition*, 1991, Bourgois/IRCAM, Paris)

connaissances et constatent que le timbre reste un « mystère »[2], un « attribut auditif mal compris »[3]. Pour eux, le concept de timbre est mal défini[4]. Il est vrai que la notion de timbre est riche et complexe et que le terme utilisé est manifestement porteur de sens multiples.

1.2. Les définitions acoustiques du timbre : le timbre et son double

Le terme « timbre » a des sens multiples et différents (voir l'encadré ci-contre). Dans le domaine de la musique, il se réfère initialement à la cloche hémisphérique frappée par un marteau et aux mélodies bien connues qui servent de support à des paroles commentant l'actualité. Dans le sens d'une caractéristique des sons musicaux, il apparaît pour la première fois dans l'*Encyclopédie*, à l'entrée « Son » et non à celle de « Timbre », ce qui mérite d'être souligné. La définition, qui est de J.- J. Rousseau, fait déjà état des difficultés qu'il y a à saisir une notion qui n'est pas un « degré » quantifiable comme la hauteur ou l'intensité, mais une « qualité » du son.

Il faudra près d'un siècle de recherches mathématiques et d'expérimentations acoustiques avant que ne soient établies les relations entre la présence de sons partiels ou d'harmoniques et les qualités du timbre, ce qui fut l'œuvre d'Helmholtz en 1863. Celui-ci établit que « le timbre de la portion musicale d'un son dépend seulement du nombre et de l'intensité des sons partiels ». Ces conclusions se retrouvent dans tous les ouvrages d'acoustique jusqu'à une époque récente. Toutefois la perception du timbre, et en particulier les relations entre les différentes compositions spectrales et les qualités des sons (qui seules concernent les musiciens), n'est qu'en partie abordée par Helmholtz (voir page 353).

Dans les années 1970, une définition du mot « timbre » est élaborée au sein des commissions de normalisation du vocabulaire de l'acoustique : respectivement l'acoustique musicale (France) et la psychoacoustique (États-Unis).

> *France 1972 – TIMBRE : caractère de la sensation auditive qui différencie deux sons de même hauteur et de même intensité et qui permet la reconnaissance de l'origine ou de l'appartenance à un groupe de sons donnés. (AFNOR[5])*
>
> *États-Unis 1973 – TIMBRE : attribut de la sensation auditive qui permet à l'auditeur de différencier deux sons de même hauteur et de même intensité et présentés de façon similaire. (ANSI [6])*

Première surprise : le timbre n'est pas défini du point de vue physique mais comme un attribut de la sensation auditive. Les rédacteurs de ces normes s'en remettent prudemment à un auditeur qui juge de la dissimilarité entre deux sons de même hauteur et de même intensité, ce qui associe de fait le concept de timbre à la catégorie des sons dits musicaux (ou périodiques), donc aux instruments de musique.

2. « Although much more is known in the mid-1990s, (...) timbre as a musical variable still remains, for the most part, a mystery... », Hajda, 1994.
3. « Timbre, in contrast to pitch and loudness, remains a poorly understood auditory attribute. » Donnadieu, 2007.
4. « Il n'existe pas de définition opérationnelle acceptée par tous les chercheurs, sur la base de laquelle pourrait se constituer un corpus de méthodes expérimentales et de modèles perceptifs. » Hajda, 1994, p. 253 (Traduction M. C.).
5. Association française de normalisation, norme NF S30-107, décembre 1972 – Vocabulaire de l'acoustique musicale. Le terme « timbre » est absent des définitions de l'acoustique générale de l'AFNOR.
6. American National Standards Institute, ANSI S3.20-1973 – Psychoacoustical terminology. Traduction J.-C. Risset, in Zenatti, p. 87. Texte original : « Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar. »

Brève histoire de la notion de timbre

Étymologie. Le mot « timbre », dont l'emploi en français remonterait au XIIe siècle (Dauzat, *Dictionnaire étymologique*, Larousse), désignait à l'origine un instrument sonore (figure ci-contre) :

Figure 7.1 Timbre sonore.
Guillemin, A., (1875), p. 109

> *« Sorte de cloche ronde qui n'a point de battant en dedans, et qui est frappée en dehors par un marteau. Le timbre d'une horloge, timbre d'un reveille-matin. » Dictionnaire de l'Académie française*, 1re éd., 1694, p. 563.

Ce terme s'appliquait également au casque (armoiries), à la tête, ainsi qu'à une grande pierre creuse servant d'abreuvoir, sans doute par analogie de forme. On note également l'emploi de « timbrer », pour imprimer une marque, une authenticité. Au XVIIIe siècle, l'acception sonore de timbre s'élargit :

> *« Il se dit quelquefois pour Le son que rend le timbre. Ce timbre est trop éclatant [autrement dit, le timbre du timbre qui est la cloche]. Il se prend quelquefois figurément pour Le son même de la voix. Et dans ce sens on dit d'une belle voix, voilà un beau timbre. Cette voix a un timbre argentin. » Dictionnaire de l'Académie française*, 4e éd., 1762, p. 834.

Jusqu'en 1723, le terme « timbre » n'apparaît dans les textes de musique que pour désigner la cloche sans marteau de la figure 7.1[1]. Les qualités des sons sont d'ailleurs rarement décrites dans les ouvrages, à l'exception des facteurs d'orgue qui ont un terme spécifique : *l'harmonie*. Dom Bédos (1770) donne cette définition : « Égaliser d'harmonie, c'est donner à tous les Tuyaux la même qualité de son et d'harmonie ; car si l'un avait le son sec & maigre, l'autre moëlleux, &c, ils ne seraient pas égalisés d'harmonie. »

Les définitions du timbre. Dans son sens moderne, le terme « timbre » apparaît pour la première fois dans l'*Encyclopédie*, à l'article « SON, en musique »[2] :

> *« La nature du son est l'objet des recherches du physicien ; le musicien l'examine seulement par ses modifications, et c'est selon cette dernière idée que nous l'envisageons dans cet article. Il y a trois choses à considérer dans le son : 1, le degré d'élévation entre le grave et l'aigu : 2, celui de véhémence entre le fort et le faible : 3, & la qualité du timbre qui est encore susceptible de comparaison du sourd à l'éclatant, ou de l'aigu [sic ;* lire : l'aigre*] au doux. »* Tome XV, p. 345, col. 1.

Dans la suite de l'article, l'auteur (J.-J. Rousseau) développe les connaissances acquises sur les deux premiers points, puis en vient au timbre :

> *« Quant à la différence qui se trouve encore entre les sons par la qualité du timbre, il est évident qu'elle ne tient ni au degré de gravité, ni même à celui de force. Un hautbois aura beau se mettre exactement à l'unisson d'une flûte, il aura beau radoucir le son au même degré, le son de la flûte aura toujours je ne sais quoi de doux et de moelleux, celui du hautbois je ne sais quoi de sec et d'aigre, qui empêchera qu'on ne puisse jamais les confondre. » Ibid.*, p. 346, col. 2.

Il est important de constater ici que le timbre n'est pas un *degré* mais une *qualité* et que bien que cet article ait trait au *Son*, le timbre est rapporté *ipso facto* à une source instrumentale connue : à la flûte, au hautbois, à la voix :

> *« Que dirons-nous des differens timbres des voix de même force et de même portée ? Chacun est juge de la variété prodigieuse qui s'y trouve. » Ibid.*, p. 346.

Le choix des timbres instrumentaux est spécifié dans les œuvres dramatiques (cantates et opéras de Montéclair à Rameau) pour créer un climat expressif bien particulier à certaines scènes (tendres, infernales) ou pour apporter une variété sonore aux différents couplets des chaconnes et passacailles. Il est à noter que le traité de chant de Bérard (1755), l'un des premiers ouvrages à décrire de façon détaillée le « beau chant », rapporte les qualités des voix à des effets expressifs[3]. Du point de vue scientifique, le timbre reste une énigme :

1. Constatation faite lors d'une recherche sur le terme « timbre » dans la thèse de Théodora Psychoyou (2003).
2. Tome XV. À noter que l'entrée *Tymbre* (sic) du tome XVI fournit toutes sortes de sens mais pas celui-ci.
3. Bérard cite les « sons violents, tendres, entrecoupés, etc. », c'est-à-dire avant tout des qualités en rapport avec l'expressivité du texte.

> *« Cependant, personne que je sache n'a encore examiné cette partie, qui peut-être, aussi-bien que les autres, se trouvera avoir ses difficultés : car la qualité du timbre ne peut dépendre ni du nombre des vibrations qui font le degré du grave à l'aigu, ni de la grandeur ou de la force de ces mêmes vibrations qui fait le degré du fort au faible. Il faudra donc trouver dans les corps sonores une troisième modification différente de ces deux, pour expliquer cette dernière propriété ; ce qui ne me paraît pas une chose trop aisée ; il faut recourir aux Principes d'acoustique de M. Diderot, si l'on veut approfondir cette matière. » Ibid.*, p. 346.

Dès 1765, le timbre est donc institué comme troisième grandeur du son en musique, sans que l'on en connaisse les causes physiques. Il est étonnant de constater que les harmoniques ne sont pas invoqués alors que Sauveur (1701) avait montré leur importance dans le Plein-jeu de l'orgue et que Rameau avait fondé sa théorie de la basse fondamentale sur la série harmonique (1722). De fait, les harmoniques expliquent la « résonance » alors que l'approche physique de la nature du timbre est toute autre. Il faut, pour l'aborder, disposer d'outils mathématiques et de moyens expérimentaux appropriés à l'analyse du son. Tout se joue dans la première moitié du XIXe siècle. En 1817, le physicien Biot émet le premier l'hypothèse du rôle possible des harmoniques dans le timbre. Les progrès sont rapides. D'une part, l'acoustique est à la mode[4] et, d'autre part, le timbre préoccupe les facteurs désireux de parfaire l'homogénéité de sonorité des instruments de l'orchestre. Lorsqu'en 1843 Berlioz publie le premier traité de l'instrumentation, il stipule dans son introduction :

> *« L'objet de cet ouvrage est donc d'abord, l'indication de l'étendue et de certaines parties essentielles du mécanisme des instruments, puis l'étude fort négligée jusquà présent, de la nature du timbre, du caractère particulier et des facultés expressives de chacun d'eux… »*

C'est Hermann von Helmholtz, physiologiste et physicien, qui fera les premières expériences décisives sur le timbre. Son ouvrage (1863), qui est rapidement traduit en français (1874) et en anglais (1885), a un retentissement considérable aussi bien dans le milieu scientifique que parmi les musiciens[5]. Malheureusement, la restriction au « timbre musical » introduite par Helmholtz sera rapidement oubliée. Pendant plus d'un siècle, le timbre reste pour nombre d'auteurs un synonyme de « contenu harmonique du son » ou « spectre »[6].

4. C'est l'époque des Chladni, Savart, Fourier, Cagniard de la Tour.
5. Voir le long article de l'entrée « Timbre » dans la première édition du *Grove Dictionary of Music* (1889), ainsi que l'entrée « Tone » du même ouvrage.
6. Citons par exemple un ouvrage d'acoustique publié en 1934 : « Le lecteur sait que le timbre d'un son dépend du nombre des harmoniques et de leurs amplitudes. » Foch, p. 114.

Mais de quels sons s'agit-il ? Est-il question de sons joués sur deux instruments de « timbre » différent comme le sont la flûte et le hautbois ou de deux sons joués sur le même instrument avec des « timbres » différents, comme le réalise un guitariste lorsqu'il change le point de pincement, ou un clarinettiste qui compare deux doigtés donnant la même note ?

1.3. Les deux écoutes du timbre : le timbre identitaire et le timbre qualitatif

Son 7.1 (10″)

Sept sons (la_3)

Son 7.2 (10″)

Flûte

Pour rendre sensible l'importance cruciale de cette question, commençons par une expérience. Écoutons successivement l'exemple **7.1** qui fait entendre la même note, la_3, jouée par sept instruments différents et l'exemple **7.2** dans lequel le son d'une note tenue se transforme de façon impressionnante. Immédiatement après l'audition, tentons d'analyser rétrospectivement comment nous avons écouté chacun de ces deux exemples de timbre. Le Son **7.1** dans lequel différents instruments se succèdent sollicite prioritairement l'identification de ceux-ci : nous écoutons la

trompette, la flûte, le violon[7]. Au contraire, le Son **7.2**, joué sur un seul instrument, nous engage à prêter attention aux changements de qualité sonore réalisés par le flûtiste (joueur de *shakuhachi* japonais). La notion de timbre est donc ambiguë puisqu'elle met en jeu deux types de perceptions qui dépendent du contexte musical mais aussi des motivations de l'auditeur.

La définition acoustique du timbre commune aux deux normes citées plus haut : « ce qui différencie deux sons de même hauteur et de même intensité », ne permet pas de lever l'ambiguïté que nous venons de mettre en évidence. Il est donc intéressant de remarquer que la définition de l'AFNOR est complétée par la phrase suivante, généralement omise dans les citations : « (et) qui permet la reconnaissance de l'origine ou de l'appartenance à un groupe de sons donnés ».

La notion de timbre recouvre donc à la fois la reconnaissance de l'instrument et l'appréciation des qualités sonores du son – double fonction bien connue des musiciens – aussi bien dans le cadre de la création musicale que dans celui de l'écoute. Par son aspect identitaire, le timbre assure une cohérence dans le suivi mélodique[8] et permet d'écouter distinctement une voix particulière au sein d'une polyphonie vocale ou instrumentale. Il arrive aussi que, par la fusion de deux « timbres », le compositeur crée une nouvelle entité sonore qui avive l'écoute des auditeurs. Toutes les musiques jouent ainsi, tantôt sur la distinction, tantôt sur la fusion des « timbres ». À un tout autre niveau, chaque musicien travaille le contrôle des paramètres de jeu pour produire des variations de sonorité à des fins expressives, des sons plus ou moins doux, brillants ou chaleureux. Ce faisant, il produit des changements de la qualité des sons, que l'on désigne aussi comme des changements de timbre.

La difficulté inhérente à l'étude acoustique du timbre tient au fait que ce concept correspond à deux écoutes totalement distinctes. Il nous faut donc envisager l'étude du timbre sous ces deux aspects et rechercher dans le son ce qui a trait, d'une part, à la reconnaissance des sources sonores et, d'autre part, à leur qualification. Il importe aussi de préciser le vocabulaire en distinguant par exemple le *timbre identitaire* et le *timbre qualitatif* en référence à ces deux façons d'appréhender le timbre. Dans la pratique musicale, c'est l'écoute des qualités des sons qui est principalement visée : l'instrumentiste attentif aux qualités de son instrument, le chef d'orchestre qui équilibre les pupitres ou le preneur de son soucieux de la qualité sonore d'ensemble travaillent le timbre qualitatif. Cependant, lorsque de nouveaux matériaux sonores sont donnés à entendre dans une œuvre, l'aspect identitaire du timbre reprend une part importante dans l'écoute. Ce que l'on désigne par « timbre » et le rôle que lui attribuent les musiciens change selon les situations, selon les musiques et leurs évolutions.

1.4. Le timbre, la musique et ses instruments

La phrase musicale du Son **7.3** est jouée sur un instrument à percussion, le tambour indien *pakhavaj*, avec lequel le musicien produit des sons différents en variant les types de frappes (voir chapitre 4, figure 4.45 pour l'analyse acoustique). Chacune d'elles est une entité acoustique complexe, au sens d'une forme temporospectrale, une entité reconnaissable que le musicien désigne d'ailleurs par une onomatopée vocale constituée d'une consonne et d'un timbre vocalique (*tin*, *daah*, *tac*, *dong*) ce qui correspond aux deux composantes du son, l'attaque et la hauteur spectrale.

Son 7.3 (17″)
Solo de pakhavaj

7. On remarquera que le son n° 2 (clarinette) et le son n°5 (saxophone) sont difficilement reconnaissables bien qu'ils soient extraits d'une base de données de sons musicaux connue pour sa qualité. Voir § 3.3.5 pour une discussion.
8. Voir chapitre 4, § 3.9.3, l'expérience de D. Wessel dans laquelle le changement d'instrumentation produit une fission de la mélodie.

Ces entités sonores étant en nombre réduit, l'écoute de la structure sonore porte essentiellement sur leur organisation en termes de durées et d'intensités.

À l'opposé, les instruments des musiques mélodiques et harmoniques sont conçus pour produire des sons de hauteur définie et, corrélativement, plusieurs instruments peuvent produire la même hauteur. Le fait de pouvoir nommer la hauteur d'un son indépendamment de l'instrument, donc de son timbre, contribue à séparer perceptivement ces deux aspects du son. C'est donc initialement dans les musiques de hauteur que le timbre conquiert une autonomie perceptive.

Au Moyen âge et à la Renaissance, les différentes parties d'une polyphonie vocale sont aussi jouées par des instruments, qu'il s'agisse d'instruments à vents ou à cordes. Le choix que l'on en fait peut cependant tenir compte du contexte culturel de la musique (musique religieuse, danse), donc de l'effet qu'ils produisent sur les auditeurs. Au XVIIe siècle, le développement de la basse continue favorise le développement de sonates pour des instruments solistes (flûte traversière, hautbois, musette, basson) qui acquièrent une identité de timbre très affirmée, ainsi qu'une grande diversité de sonorités selon les notes jouées.

Aux XVIIIe et XIXe siècles, la fonction du timbre dans la musique se modifie avec le développement de l'orchestre, véritable laboratoire du matériau sonore. Les instruments se transforment pour s'adapter aux nouvelles exigences sonores de l'écriture musicale qui sont, au premier chef, l'homogénéité de sonorité d'un type d'instrument donné sur toute la tessiture et, conjointement, le besoin d'unir les instruments en masses d'intensités et de sonorités différentes. Les compositeurs développent une écriture combinant timbre et hauteur en jouant sur les oppositions de sonorités, sur les doublures, voire en alternant les instruments au sein d'une même mélodie[9]. En marge de l'orchestre, le quatuor à cordes est aussi une formation pilote par laquelle l'écriture musicale innove et expérimente sans cesse les effets de fusion ou d'individuation des instruments, les inversions de tessiture et les nouveaux modes de jeu.

1.5. Les nouvelles écritures orchestrales

> *Je le répète : le son est le matériau de la musique. Il doit donc, en vertu même de toutes ses propriétés et des effets qui en découlent, être considéré comme détenteur d'un pouvoir artistique.*
>
> Schoenberg, A., 1911, p. 39.

Au début du XXe siècle, un grand nombre de compositeurs souhaitent trouver d'autres modes d'organisation du son et travaillent en particulier à une transformation de la matière sonore de l'orchestre par le biais de l'écriture.

Hauteur ou timbre

Bartók, Berg, Schoenberg usent de groupes de sons simultanés dont la complexité est telle que la perception est moins celle d'un accord, à proprement parler, que d'une couleur sonore (*Klangfarbe*). Dans le dernier chapitre de son *Traité d'harmonie*[10], Schoenberg parle de ces « nouveaux agrégats pleins et luxuriants » et ajoute : « Quoi qu'il en soit, l'attention que nous portons aux timbres est de plus en plus vive. » L'enchaînement d'agrégats positionnés différemment dans la tessiture fait naître des mélodies spectrales : « La couleur du son est donc le grand territoire dont une région est constituée par la hauteur du son. » Pourtant, son rêve d'une organisation des timbres, à l'instar des hauteurs, se heurte à l'impossible hiérarchie des timbres.

9. Un bel exemple en est la « Marche au supplice » (mesures 82 à 87) de *La Symphonie fantastique* de Berlioz.
10. Page 515 de l'édition traduite en français.

Debussy (*Jeux,* 1913) et Schoenberg (*Farben,* 1912) sont cités comme les deux premiers compositeurs ayant orienté leur écriture musicale de façon à stimuler une écoute globale de la couleur orchestrale dans laquelle les instruments individuels ne sont plus reconnaissables en tant que tels.

Neutralisation de la hauteur

Pour beaucoup de compositeurs travaillant à structurer leur musique sur la base du timbre, la recherche passe par l'abandon d'une écriture des hauteurs et une exploration de la matière sonore. La première étape consiste donc à « brouiller » la hauteur musicale :

1. par accumulation de « notes » différentes (agrégats sonores) ;
2. par accumulation de hauteurs voisines simultanées (micropolyphonie de 48 parties et plus dans *Atmosphères* de Ligeti, 1961) au profit de microvariations indiscernables produisant le *timbre de mouvement*[11] ;
3. par l'emploi de percussions de hauteur indéterminée. Varèse (*Ionisation* pour 37 percussions, 1931), qui en est un des pionniers, n'a cessé de dire qu'il écrit « le son ». Dans plusieurs de ses textes, Varèse préfigure les développements techniques qui vont permettre de travailler directement sur le son, voire de le créer[12].

Quand la hauteur est présente, elle est figée et statique de sorte que l'auditeur se concentre sur les autres variations des sons : sur les battements et les frottements entre hauteurs voisines (*Quatre Pezzi* sur une seule note, Scelsi, 1959), sur les variations d'intensité des harmoniques et les trajets spectraux des formants de la voix (*Stimmung* de Stockhausen, 1968), plus tardivement sur une véritable synthèse instrumentale dans les musiques dites spectrales (*Partiels* de Grisey, 1975) ;

1.6. Les nouvelles machines sonores

Parallèlement aux recherches développées par les compositeurs attachés aux instruments de l'orchestre apparaissent des instruments d'un type nouveau issus des possibilités sonores offertes par les technologies du téléphone (Cahill, vers 1900), de la radio (vers 1920, le Theremin puis les ondes Martenot), de l'électronique (synthèse additive, vers 1936, l'orgue Hammond et vers 1960, l'orgue Dereux électrostatique) ; de la synthèse numérique par Voltage Control (Moog, années 1970), par modulation de fréquence (DX7 Yamaha, années 1980). Les sons « inouïs » de ces instruments intriguent, provoquent l'étonnement ou l'agacement des auditeurs. Mais, la plupart ayant été conçus pour produire des sons de hauteur définie, ils s'intègrent rapidement aux formations instrumentales traditionnelles tout en constituant une nouvelle catégorie, celle des sons électroniques ou, plus généralement, des sons de synthèse[13].

Une étape beaucoup plus radicale se prépare. Elle conjugue l'attrait pour les sons des machines et, d'une façon générale, pour toutes sortes de sons, étrangers à ceux de l'orchestre. Dès 1913, les bruiteurs futuristes clament qu'il faut « remplacer la variété restreinte des timbres des instruments que possède l'orchestre par la variété infinie des timbres des bruits » (Russolo, 1916). Ils construisent alors des instruments mécaniques, censés reproduire et jouer les bruits de notre environnement, qui portent les noms évocateurs de glouglouteurs, crépiteurs, froufrouteurs, et hululeurs. Leurs concerts, souvent hués, marquent les esprits[14].

11. Voir Ligeti, 2001, chapitre IX, p. 198-209.
12. Voir Varèse, 1983, p. 62-63, textes de 1930.
13. « Listeners tend to hear electronic sounds as a class – ah, electronic – just as they might hear – *ah, a violin* », Erickson, (1975), p. 8.
14. Voir aussi John Cage. Pour son ballet *Bacchanale* (1938), le compositeur désirait utiliser un orchestre de percussions mais n'en disposait pas. Il utilisa donc le piano présent dans la salle en garnissant les cordes d'objets divers : c'est le premier *piano préparé*. Voir Boulez-Cage, 2000, p. 72.

Varèse, qui est depuis longtemps en quête de nouveaux sons, proteste à sa manière :

> *Pourquoi, futuristes italiens, reproduisez-vous servilement la trépidation de notre vie quotidienne en ce qu'elle n'a que de superficiel et de gênant ? Je rêve les instruments obéissants à la pensée – et qui avec l'apport d'une floraison de timbres insoupçonnés se prêtent aux combinaisons qu'il me plaira de leur imposer et se plient à l'exigence de mon rythme intérieur.*
>
> Varèse, E., juin 1917 cité dans *Écrits*, 1983, Ch. Bourgois, p. 143.

Dans les années 1950, l'électronique, l'enregistrement et la reproduction des sons lèvent toutes les contraintes. Il devient possible de s'affranchir des instruments habituels et d'accéder directement à « l'océan des sons » dont rêvait Jörg Mager, en façonnant directement le matériau brut de la musique. La naissance de studios de musique expérimentale[15], qui rassemblent musiciens et chercheurs scientifiques, s'avère déterminante en raison de la confrontation permanente qui s'opère entre l'expérimentation sonore et l'écoute[16]. À l'opposé d'un ingénieur qui crée des machines dont il découvre les sons, un musicien est entraîné à « entendre » intérieurement le son avant de le produire. En s'appropriant les possibilités de la synthèse électronique, des compositeurs comme Stockhausen ou Ligeti questionnent la perception sonore et consignent leurs réflexions, et leurs déceptions, dans des textes importants[17]. À Paris, l'orientation prise par Pierre Schaeffer et Pierre Henry, qui privilégie les transformations de sons enregistrés, est fertile en réflexions sur les deux écoutes du son qui nous préoccupent.

1.7. Les écoutes du matériau sonore : Pierre Schaeffer et les musiques électroacoustiques

> *Oubliant délibérément toute référence à ces causes instrumentales ou à des significations musicales préexistantes, nous cherchons alors à nous consacrer entièrement et exclusivement à l'écoute, à surprendre ainsi les cheminements instinctifs qui mènent du pur « sonore » au pur « musical ». Telle est la suggestion de l'acousmatique : nier l'instrument et le conditionnement culturel, mettre face à nous le sonore et son « possible » musical.*
>
> Schaeffer, P., 1966, p. 98.

Nous sommes tellement habitués aujourd'hui à entendre les musiques les plus diverses que nous avons peine à imaginer le choc ressenti par les premiers auditeurs d'un concert de bruitistes ou d'un concert de musique concrète. Il faut lire les premières pages du journal que Schaeffer a tenu dans les premiers mois de son aventure au Service de la recherche de la radio pour mesurer l'importance des chocs perceptifs déclenchés par les premières expériences de manipulation du son que permettait le travail en studio.

En prélevant des sons dans l'univers sonore quotidien (train, casseroles, tourniquet), Schaeffer s'est heurté de front à l'écoute causale, celle qu'il appelle « anecdotique » et qui, en quelque sorte, fait écran à l'écoute musicale. « Le bruit des tampons[18] n'est-il pas d'abord anecdotique, donc antimusical ? » (1952, p. 20). À la même époque, l'expérience fortuite du « sillon fermé » met Schaeffer sur la voie d'un nouveau type d'écoute. Un fragment de son de cloche, prisonnier d'un sillon de disque fermé que l'appareil lit en boucle, est rendu méconnaissable par

15. Studios de la RTF à Paris (P. Schaeffer, P. Henry, A. Moles), de la RAI à Milan (L. Berio, B. Maderna, 1954) de la WDR à Cologne (H. Eimert, G. Ligeti, K. Stockhausen), Studio de Gravesano (H. Scherchen, 1954). Voir le site http://sonhors.free.fr/.
16. Voir à ce sujet le remarquable ouvrage d'A. Moles, *Les musiques expérimentales*, 1960.
17. Stockhausen, 1988 ; Ligeti, 2001.
18. Bruit de chocs que font en se heurtant les « tampons » des wagons de chemin de fer.

l'opération[19] et provoque une écoute « décalée ». Ce n'est plus la cloche qu'il écoute mais le son pour lui-même, la matière sonore constitutive du son de cloche initial. Pour créer une musique hors du champ instrumental habituel, Pierre Schaeffer tire alors la conclusion suivante :

> *Deux démarches sont préalables : distinguer un élément (l'entendre en soi, pour sa texture, sa matière, sa couleur) ; le répéter. Répétez deux fois le même fragment sonore : il n'y a plus événement, il y a musique.*
>
> Schaeffer, P., 1952, p. 21.

En d'autres termes, Pierre Schaeffer enjoint l'auditeur à s'affranchir de l'écoute événementielle (causale, anecdotique) au profit d'une écoute dédiée au matériau sonore pour lui-même, ce qu'il nommera par la suite l'« écoute réduite ». Que devient alors le timbre dans cette nouvelle musique ?

> *Il n'y a pas d'instrument à jouer de la musique concrète. Telle est la difficulté majeure.*
>
> Schaeffer, P., 1952, p. 26.

L'abandon des instruments de musique façonnés au cours des siècles pour produire des sons *aptes au musical*, n'est pas sans problème. Les sons créés en laboratoire n'ayant pas d'identité, comment les désigner ? Que faut-il écouter ? Confronté à ces questions, Pierre Schaeffer définit le concept « d'objet sonore » et entreprend, avec les musiciens réunis au Club d'essai de la RTF, un inventaire des traits morphologiques acoustiques permettant de situer chaque objet sonore dans l'espace d'une nouvelle cartographie sonore. L'entreprise est immense et, surtout, nécessite une formation à la nouvelle écoute dite acousmatique, que les auditeurs devront apprendre à partager. Ces problématiques sont abondamment décrites et débattues dans les ouvrages du Groupe de recherches musicales (GRM)[20]. Les musiques électroacoustiques, qui mettent en jeu de complexes manipulations spectrales et une grande diversité d'articulations spatio-temporelles de la matière sonore, appellent de nouvelles méthodes d'analyse musicologique. Stéphane Roy[21] est l'auteur d'une thèse dans laquelle il compare différentes méthodes d'analyse des musiques électroacoustiques. Il souligne l'importance perceptive de deux critères : le critère de signature timbrale ou d'identité timbrale des unités et les critères de variations (dynamique, hauteur, durée, timbre et registre) de ces unités. En conclusion, il avance que la discrimination fine des variations, des petits écarts qui font toute la richesse de l'écoute qualitative, doit, pour être ressentie, s'appliquer à des entités repérables par leur identité timbrale. C'est un joli retour du « timbre identitaire », que l'auditeur doit recréer afin d'entrer dans ce nouvel univers sans instrument.

1.8. La synthèse sonore, pierre de touche de la perception

Lorsque la synthèse sonore numérique autorise enfin, dans les années 1980, un contrôle fin et précis de tous les paramètres du son[22], de jeunes compositeurs désireux de façonner eux-mêmes leurs sons mettent leur espoir dans un renouvellement de l'écriture musicale du timbre. Le travail est souvent laborieux, et les déceptions qui s'expriment dans les quelques citations qui suivent dénotent l'impossible saisie de l'objet même de la synthèse : le timbre !

19. À la vitesse de 78 tours par minute, une boucle sonore ne dure que 760 ms. De plus, le fragment est privé de l'attaque de la cloche.
20. En particulier l'ouvrage collectif *Du sonore au musical, cinquante années de recherches concrètes (1948-1958)*, Veitl A., Dallet S., 2001.
21. *L'analyse des musiques électroacoustiques : modèles et propositions*, 2003, l'Harmattan, p. 285.
22. Voir à ce sujet J.-C. Risset, *Timbre*, 2004, p. 87-114.

Lorsqu'on crée des sons de synthèse en partant de zéro, on constate très vite que le timbre existe à peine en tant que concept.

Harvey, J., 1991, collectif-Barrière, p. 456.

Si j'en reviens à la notion de timbre, instrumental ou vocal, le timbre existe esthétiquement lorsqu'il est directement relié à la constitution proprement dite de l'objet musical. Le timbre en soi n'est rien, comme un son n'est rien.

Boulez, P., 1991, *ibid.*, p. 545.

Ce que nous appelons timbre est une donnée perceptuelle et non conceptuelle. Il touche un domaine de la perception qui est essentiellement synthétique et immédiat. On peut le créer dans certains cas bien particuliers de fusion entre divers composants de base mais, dégagé de toute référence à un modèle connu ou mémorisé, il est une donnée stérile de la composition.

Manoury, Ph., 1991, *ibid.*, p. 300.

Le constat général est que le timbre ne peut être dissocié de la hauteur, de l'espace et de la texture harmonique. Sa prise en compte dans l'écriture s'inscrit toujours dans un projet esthétique donné, lequel conditionne la façon dont sonne la musique et la façon dont elle agit sur les auditeurs, ce que J.-C. Risset résume dans une élégante formule :

La musique est flux, mouvement, elle n'est pas une forme abstraite, elle n'existe que par son incarnation dans le temps et dans le son : le timbre est sa substance même.

Risset, J.-C., 1991, *ibid.*, p. 239.

1.9. Du timbre au « son pour lui-même » : vers une dissolution du concept de timbre ?

Timbre de la musique, timbre d'un instrument, timbre d'un son : plus encore que pour les notions de hauteur et d'intensité le timbre dépend de notre écoute. Sous ses deux aspects causal et qualitatif, il est la modalité perceptive par laquelle nous appréhendons la matière sonore de la musique, celle dans laquelle s'incarnent la hauteur et l'intensité et relève des processus de catégorisation, notamment celui de l'identification sonore, *Quoi ? Où ?* Ces questions superflues à la vue d'un orchestre classique, s'invitent à l'écoute des musiques sur support qui organisent des sons nouveaux (bruits et sons de synthèse), ou des sons de l'environnement. S'efforçant de gommer les traits anecdotiques des sons, les compositeurs enjoignent l'auditeur à explorer de nouveaux territoires afin de développer une écoute du *timbre-couleur* pour construire de nouvelles catégories de qualités sonores. L'attention auditive change d'objet et passe de l'écoute de sons instrumentaux parés de qualités sonores variées à celle du *son pour lui-même*, un son qui n'est plus rapporté à un mode de production connu[23]. Lorsque l'instrument a disparu, lorsque le son est façonné en même temps que la musique, la notion même de timbre devient difficile à appréhender.

L'étude acoustique et perceptive que nous proposons dans ce chapitre s'inscrit dans l'univers instrumental de l'Occident, des réalisations mécaniques aux productions sonores numériques.

23. Comme en témoigne le récent ouvrage de Makis Solomos : *De la musique au son : l'émergence du son dans la musique des XXe-XXIe siècles*, Presses Universitaires de Rennes (546 p.).

2. Le timbre identitaire des sources instrumentales

Le matériau brut de la musique est le son.

Varèse, E., 1939, *in* Charbonnier, 1970, p. 91.

Nous venons de voir que ce qu'on désigne par « timbre » change de sens selon les types d'écoutes et selon les musiques. Nous avons remarqué aussi que les différentes fonctions du timbre analysées par les compositeurs, les musicologues et les instrumentistes peuvent toujours se ramener aux deux modalités fondamentales de la perception sonore analysées au chapitre 4, qui sont l'identification des sons et leur qualification. En tant que matière sonore de la musique le timbre est tantôt le marqueur signalétique d'une identité sonore (voix ou instrument) et tantôt l'indicateur d'effets produits par le son d'un instrument, d'un agrégat d'instruments ou encore de sons enregistrés et travaillés en studio. Compte tenu des interactions complexes observées entre sons, musiques, timbres et perception, nous proposons d'explorer initialement l'aspect proprement causal du matériau sonore, à l'aide d'un inventaire acoustique des sources sonores typiques, nous réservant d'aborder ultérieurement (voir page 346) les aspects qualitatifs, qui dépendent pour une grande part de la culture des auditeurs et de leur expertise d'écoute.

2.1. La typologie acoustique des formes sonores de base

2.1.1. La forme spectrotemporelle

La perception d'un son déclenche, consciemment ou non, la question : « que s'est-il passé ? » – ce qui en musique se traduit par : « quel est cet instrument ? ». Y répondre, c'est être capable de repérer la forme sonore entendue ; tout se passe très vite lorsqu'il s'agit d'un son connu. Au niveau de base d'une catégorisation sonore, celui auquel la plupart des auditeurs reconnaissent sans peine les grandes familles de sons musicaux (flûte, piano, trompette, tambour et cymbales), la catégorisation perceptive des sons de la musique ne diffère pas de celle des sons de l'environnement (voir chapitre 4, § 3.4) : elle s'appuie sur la forme spectrotemporelle des sons qui est la combinaison de deux critères.

Critère 1 – Dynamique temporelle. Le premier critère a trait à l'action par laquelle l'énergie est communiquée à une structure quelconque : choc unique ou geste continu (figure 7.2a). Cette dualité fondamentale (voir chapitre 1, § 2.2), qu'André Souris (p. 83) rapproche de l'opposition entre le jeu *détaché* et le *legato*, distingue les deux grandes catégories instrumentales : d'un côté les instruments dans lesquels le son est produit par une impulsion (percussions, cordes frappées et pincées) et de l'autre ceux dans lesquels le son est entretenu (vents et cordes frottées). Du point de vue perceptif, les sons obtenus par percussion se signalent par une caractéristique importante : immédiatement après le maximum de l'impact, l'énergie décroît jusqu'à l'extinction, ce qui produit une enveloppe dynamique asymétrique, non réversible temporellement[24]. Rares sont les instruments comme le steel-drum (voir chapitre 2, § 4) qui jettent le trouble dans cette dichotomie. Entre l'impact isolé et l'entretien continu se place le groupe des sons produits par itération d'un impact à une cadence comprise entre 10 et 20 Hz : roulements, raclements et grincements par friction. Ces modes de jeu permettent de prolonger la durée de la vibration mais relèvent toujours de la catégorie des sons impulsifs. Ils possèdent une enveloppe dynamique reconnaissable à son caractère « rugueux », voire agressif (par exemple, le *guiro*).

24. L'inversion temporelle correspond à un son d'intensité croissante qui se termine brutalement à son maximum ; l'effet surprend toujours (voir chapitre 4, § 3.5).

Critère 2 – Distribution spectrale. Le deuxième critère concerne la répartition de l'énergie dans le spectre (figure 7.2b). Les premiers instants d'un son portent également des indications sur la façon dont l'énergie se développe dans le spectre des fréquences en fonction des modes vibratoires de la structure excitée. Contrairement à la dichotomie temporelle précédente, tous les intermédiaires existent entre différentes allures spectrales : spectre continu (bruit blanc) ou coloré, spectre de partiels, spectre harmonique, spectre constitué d'une seule fréquence (rare dans la réalité instrumentale)[25].

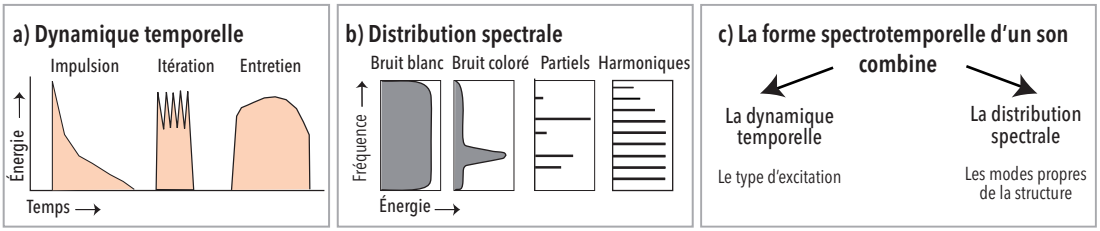


Figure 7.2 Typologie schématique de catégories temporelles et spectrales fondamentales.

La combinaison acoustique de la dynamique temporelle et du contenu spectral produit une forme spectrotemporelle (figure 7.2c). On peut ainsi établir une typologie acoustique des sources usuelles au niveau de base de la catégorisation sonore. Parmi tous les sons possibles, ceux que développent les musiciens possèdent une forme spectrotemporelle « forte » (voir chapitre 4, § 3.2), caractérisée principalement par une signature spectrale (bruit coloré, spectre harmonique). La position de cette signature selon la dimension des fréquences revêt donc une importance particulière dans l'écoute musicale, tant pour la reconnaissance que pour l'appréciation des qualités des sons.

Tessiture, hauteur spectrale

2.1.2. Quelques particularités de la dimension spectrale

La dimension spectrale correspond à l'espace des fréquences audibles, qui s'étend schématiquement de 20 à 20000 Hz. Le déplacement de l'énergie le long de cet axe produit des effets perceptifs assez complexes. Si nous demandons à quelques auditeurs sans formation musicale particulière de réagir à l'écoute de deux sons enregistrés, par exemple un coup de grosse caisse suivi d'un coup de caisse claire, nous pouvons penser recueillir les réponses suivantes : l'un dit qu'il a entendu deux tambours, un gros et un petit ; un autre qu'il a entendu « poum, tin » ; un troisième dit que le premier coup est sourd et grave, et le deuxième plus clair ou plus aigu. Par ce seul exemple se manifestent déjà diverses interprétations perceptives de la dimension spectrale : une inférence sur les dimensions de l'objet excité (gros/petit), une interprétation en termes de « hauteur » (grave/aigu), une analyse de « qualité » sonore (sourd/clair), une transposition en langue rendue par le contraste entre deux onomatopées (poum, tin).

L'association entre les dimensions d'un objet et la tessiture, grave ou aiguë, des sons qu'il émet est une donnée mécanique d'expérience courante. Les gros objets sont généralement plus lourds et produisent des sons plus graves que les petits. La hauteur spectrale qui correspond au maximum d'énergie dans le spectre, et dont le codage neurophysiologique est distinct de celui de la hauteur tonale ou fondamentale (voir chapitre 3, § 1.5), est un important critère de discrimination sonore. De récentes études ont montré que, dans certaines sociétés de primates, les résonances du conduit

25. Rappelons toutefois que le spectre harmonique ne se rencontre que dans la catégorie des sons entretenus.

vocal sont utilisées comme indice de la taille d'un individu, indépendamment de la hauteur fondamentale de la voix (voir Fitch, 2006 ; Patterson *& al.*, 2010).

Les liens entre tessiture spectrale et qualités perçues sont plus complexes. Ils font intervenir la sensibilité de l'oreille qui varie selon la fréquence (les sons aigus sont stridents, perçants) ainsi que ses capacités de discrimination temporelle (les sons graves sont plus confus). Par ailleurs, la hauteur spectrale peut interférer de façon complexe avec la hauteur tonale, au point que comparant deux sons harmoniques de même fréquence fondamentale, celui qui a le plus d'harmoniques paraît souvent plus clair et même plus « haut » que les autres (voir chapitre 6, § 2.4).

L'imitation par onomatopée est tout à fait intéressante. Avec la bouche, nous pouvons produire un éventail sonore d'une grande diversité : des clics et des clocs, des chuintements sourds ou stridents, des grognements et des sons périodiques (voir chapitre 2, § 2), toute une richesse sonore exploitée par les artistes du *beatbox*. Lorsqu'elle est bien faite, l'imitation donne ainsi de précieuses indications sur la perception car, dans sa reproduction, l'imitateur ne conserve que les traits nécessaires et pertinents à l'écoute. L'imitation de trois sons de percussion de hauteur grave, médium et aiguë, que l'on peut faire en articulant « poum, tan, tin[26] » – de préférence en voix chuchotée pour éliminer la hauteur fondamentale donnée par les cordes vocales –, fait apparaître sur l'analyse spectrale de la figure 7.3 le mouvement spectral ascendant dû au deuxième formant vocalique. En musique tibétaine, les sons joués par les grandes trompes « dung-chen » sont classés en trois groupes : grave, médium, aigu (voir Helffer, 1994, p. 38). Les termes qui les désignent, bien qu'ils varient selon les écoles, comportent tous la voyelle « o » pour les sons graves, la voyelle « a » pour le médium et la voyelle « i » pour les sons de l'aigu.

D'une façon générale, lorsqu'un son instrumental possède une signature spectrale formantique située dans la zone 400-2 500 Hz, la hauteur spectrale prend un sens vocalique. Cependant, les sons vocaux induisent un mode d'écoute particulier qui est celui des catégories phonétiques d'une langue. Pour certains auditeurs, l'effet est si net qu'une sinusoïde glissante du grave à l'aigu leur donne l'impression de changer de « timbre », de passer par le /ou/, le /o/, etc. jusqu'au /i/[27] alors même que le spectre d'un tel son est réduit à une seule fréquence.

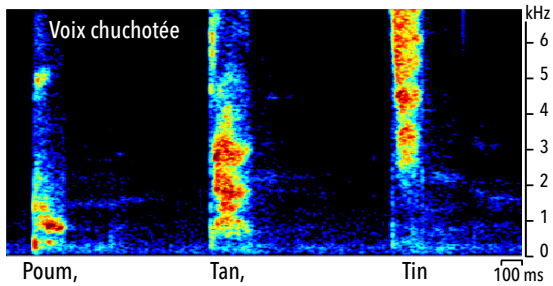


Figure 7.3 Sonagramme de « poum, tan, tin » dit en voix chuchotée. Voix féminine.

En résumé, la dimension spectrale joue un rôle primordial dans la catégorisation des sons. Elle participe aussi bien de la reconnaissance des formes sonores que de leur qualification. Toutefois, les effets perceptifs qu'elle induit s'interprètent différemment selon qu'il s'agit de sons possédant ou non une hauteur précise.

Après ce rappel des ingrédients acoustiques de base entrant dans les catégories sonores les plus générales, voyons comment se déclinent celles des sources instrumentales, en commençant par les instruments dans lesquels le son est produit par une ou plusieurs impulsions.

26. Prononcer « ta(ne), ti(ne) », avec un « e » muet.
27. Voir Ernst Mach, 1885, article cité par R. Cogan, 1984, p. 12, note 9.

2.2. Typologie des sources sonores impulsionnelles (percussions)

Un morceau de bois donne d'abord un timbre.

Souris, A., 1976, p. 97.

Du point de vue acoustique, cette catégorie comprend les instruments dits « à percussion », ainsi que les instruments à cordes non entretenus : les cordes pincées et frappées. Il existe un nombre incalculable d'instruments à percussion. Dans la musique occidentale, la plupart des ouvrages s'accordent pour distinguer trois grandes catégories : les timbales, les claviers et les objets qui forment le groupe des instruments « à sons indéterminés »[28].

2.2.1. Les percussions de hauteur indéterminée

Les instruments de percussion de hauteur indéterminée forment un groupe au sein duquel la distinction se fait sur la base du matériau mis en vibration. Indépendamment du corps humain, largement mis à contribution dans de nombreuses cultures musicales, tous les matériaux peuvent servir à la fabrication d'objets à percuter, mais ce sont principalement le bois, la peau et le métal qui sont le plus utilisés dans l'orchestre. Ces trois types de matériaux diffèrent par leur coefficient d'amortissement. De cette caractéristique mécanique dépendent à la fois la durée de la vibration et l'allure générale du spectre de fréquences – en un mot la forme spectro-temporelle. La durée de vibration est donc un paramètre fondamental des percussions.

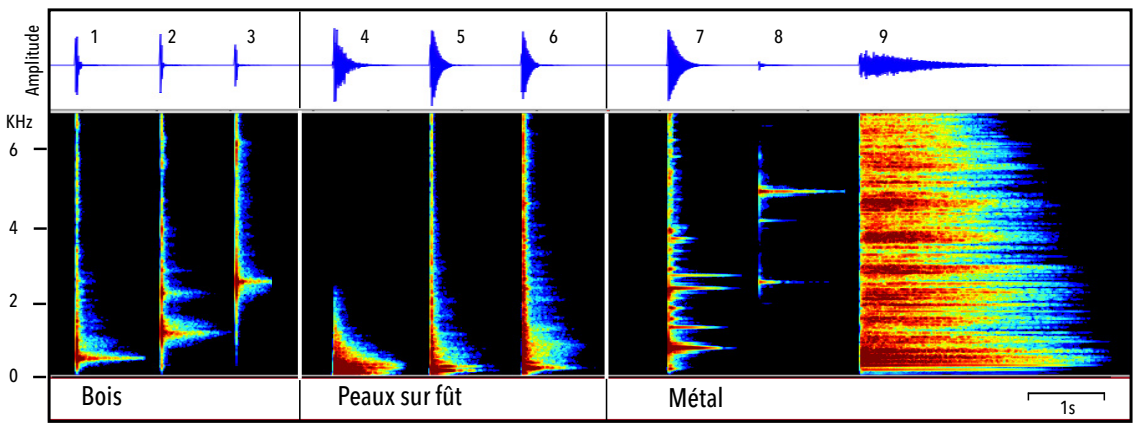


Figure 7.4 Analyses spectrales de quelques exemples de sons (du grave à l'aigu), produits par les trois groupes d'instruments à sons indéterminés. BOIS : 1. *temple block* ; 2. *wood block* ; 3. clave. PEAUX : 4. grosse caisse ; 5. tumba ; 6. tom-tom. MÉTAL : 7. cloche à vache ; 8. triangle ; 9. cymbale crash.

Lorsque le choc est précis et très bref, comme avec le bois, l'instrument peut produire des rythmes rapides et des répétitions très distinctes (castagnettes), mais la hauteur spectrale est imprécise. Les peaux, dont la durée de vibration est un peu plus longue, offrent une grande diversité de sonorités par le choix du point de frappe et la durée du contact, surtout lorsque le jeu se fait à mains nues. Le métal, qui a un faible amortissement favorise la production de modes propres de fréquences bien définies ; pendant la durée de la résonance, la vie spectrale est d'une grande complexité (cymbales, gongs et cloches). Pour tous les instruments à percussion, la dureté du percuteur, sa surface de contact et son poids jouent un rôle important dans la sonorité. Quelques échantillons typiques de ces trois groupes

28. Ainsi désignés parce qu'on ne peut pas leur attribuer une hauteur précise assimilable à une « note ».

d'instruments sont rassemblés dans l'exemple sonore **7.4** et leurs formes spectro-temporelles sont analysées figure 7.4.

Son 7.4 (16″)

Percussions : bois, peaux, métal

À l'intérieur des trois groupes d'instruments à sons indéterminés se déclinent différentes signatures spectrales et, pour chaque type d'instrument, des transpositions d'échelle en tessiture en fonction de leurs dimensions. On rencontre couramment trois tailles : grave, medium, aigu. Quelques instruments ont subi des adaptations par association de cavités (*temple block*) ou traitement des membranes (tambours de l'Inde et de l'Asie) afin d'émettre une hauteur repérable.

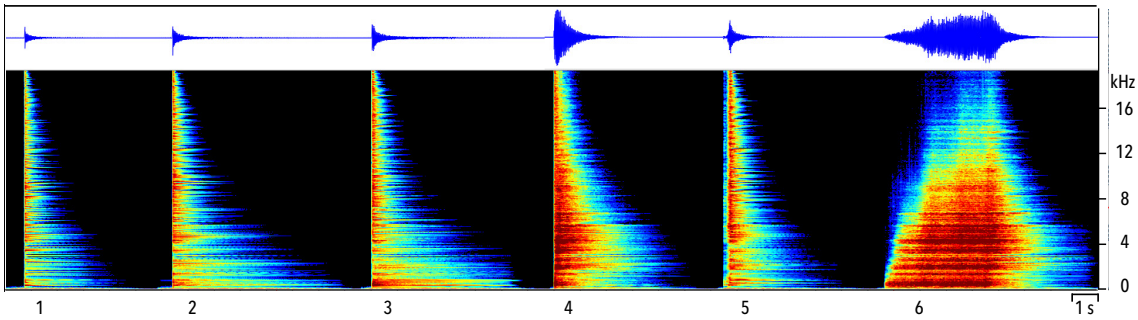


Figure 7.5 Une cymbale et ses différents modes de jeu. Successivement : 1. frappe au centre ; 2. au milieu ; 3. au bord ; 4. crash *mf* ; 5. glissé avec la brosse métal ; 6. roulement *crescendo* avec baguettes douces.

A priori, chaque objet produit une forme spectrale définie par construction (dimensions, matières) que le musicien ne peut modifier que de façon limitée puisque le contenu spectral est « donné » par les modes propres de l'objet percuté[29], ce que contestent les percussionnistes expérimentés. On peut modifier la qualité des sons par le choix des percuteurs (mou, dur ou multi-impact), en variant le point d'excitation, la force et les types d'impact, en jouant sur la durée du contact, l'étouffement et les redoublements (voir J.-C. François, 1991, chapitre 2). Certains objets comme les cymbales, riches en partiels, dont la réponse spectrale varie largement avec le point de frappe et avec l'accumulation d'énergie, deviennent, sous les doigts d'un musicien qui les maîtrise, des instruments d'une variété sonore inépuisable (Son **7.5** et figure 7.5).

Son 7.5 (37″)

Cymbale : six modes de jeu

Plusieurs instruments issus des cultures du monde ont été introduits dans l'instrumentarium en apportant une grande diversité d'effets, intermédiaires entre le choc unitaire et les rythmes de la batterie classique. On en joue par entrechocs (objets suspendus en bois ou en métal), par raclements (courge striée), secouements (hochets emplis de graines sèches), c'est-à-dire au moyen de gestes permettant de créer des « textures sonores » continues, dont la rugosité varie selon la rapidité des mouvements.

Leurs sons couvrent un large spectre, principalement dans l'aigu ou le suraigu, ce qui, combiné au bruit du choc, leur confère une bonne émergence perceptive sur un ensemble instrumental (Son **7.6**). Notez que l'échelle en fréquence de l'analyse spectrographique de la figure 7.6 couvre toute la zone audible (jusqu'à 20 kHz).

Son 7.6 (8″)

Instruments de texture rythmique et spectrale

29. « Sur ce matériau, le compositeur ne dispose d'aucun pouvoir, il se borne à prendre les instruments qui lui dont donnés. » (Boulez, 1985)

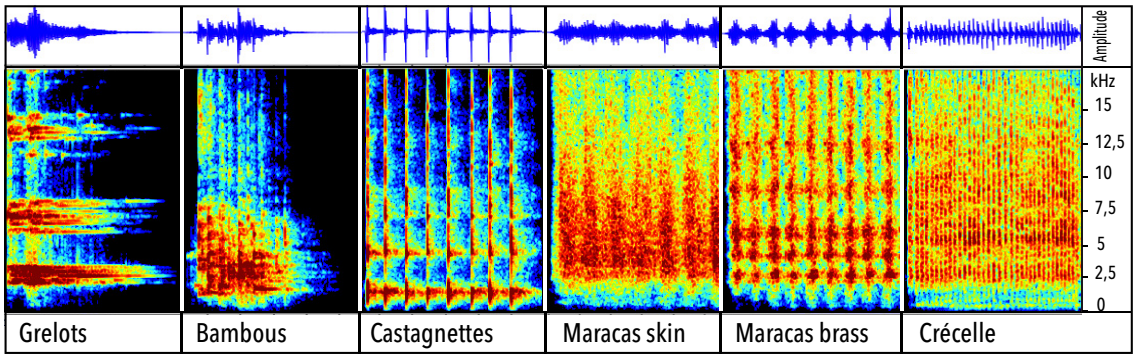


Figure 7.6 Textures rythmiques et spectrales d'instruments à secouement, roulement, raclement. Amplitude (haut de la figure) et sonagramme. Chaque extrait dure 0,55 s. Échelle de fréquences : 0 à 17 kHz ; FFT : 512 pts. Notez que l'énergie se développe au-dessus de 1 kHz.

Son 7.6 (8")

Instruments de texture rythmique et spectrale

Son 7.7 (10")

À l'extrême diversité de sonorités des instruments à percussion sans hauteur déterminée correspond une grande richesse de formes acoustiques identitaires dont la représentation sonagraphique rend compte avec bonheur (Son **7.7** et figure 7.7).

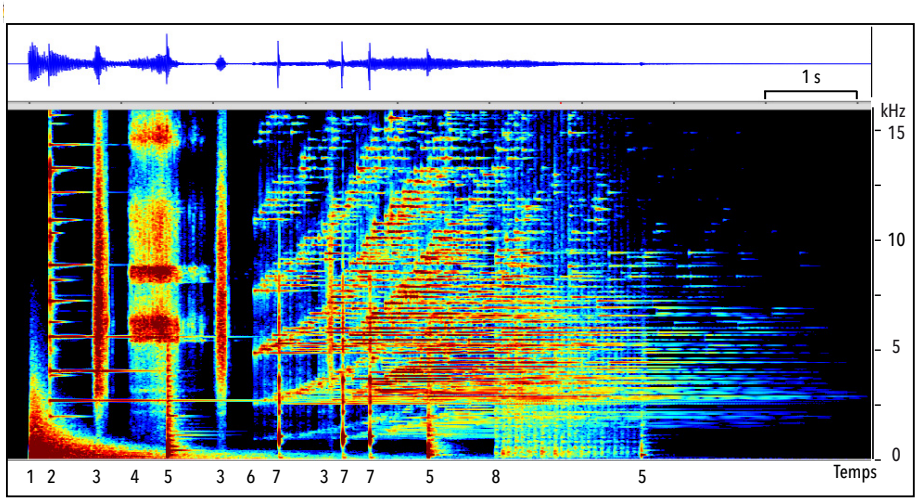


Figure 7.7 Séquence musicale d'instruments à percussion de hauteur indéterminée (voir chapitre 2, § 4.12 pour le détail des instruments).

2.2.2. Les percussions de hauteur déterminée : la « note »

Les timbales et les claviers apportent un important changement dans le concept d'instruments à percussion et surtout dans la perception des sons qu'ils émettent. La cymbale (ou de façon analogue un *pakhavaj* indien) est un objet unique sur lequel le musicien s'ingénie à varier les modes de jeux. À l'opposé, les claviers et les timbales sont des structures rassemblant chacune une série d'objets similaires par leur forme et leur matériau, qui ne diffèrent que par la « note » à émettre. La hauteur mélodique devient alors la variable d'écoute principale au point que le « timbre », commun à ces objets, passe au second plan. En d'autres termes, dans la catégorie instrumentale des percussions de hauteur déterminée, hauteur et timbre causal sont dissociés.

Toutefois, la suite des partiels d'une lame, d'une membrane ou d'un tube rigide s'écartant passablement de la série harmonique, et cela de façon différente pour chaque type d'instrument, la hauteur perçue reste tributaire du contenu spectral des sons rayonnés, et donc du timbre causal. Elle peut aussi se modifier avec la tessiture de l'instrument (voir chapitre 6, § 3 et 4). Écoutons les exemples sonores suivants dans lesquels la même note est jouée sur différents instruments à percussions, d'abord sol_2 (Son **7.8**) puis sol_3 (Son **7.9**).

Son 7.8 (9″)

Note sol$_2$, trois instruments

Son 7.9 (19″)

Note sol$_3$, six instruments

a/ Tubes, lames et membranes

Deux instruments se font remarquer à l'écoute : la timbale, dont la hauteur paraît beaucoup plus grave que le marimba qui lui succède, et la cloche tube dont la « note » est difficile à appréhender après celle du vibraphone[30]. Nous avons analysé en détail la perception de hauteur de ces deux instruments au chapitre 6 (voir § 3.4). La figure 7.8 montre l'analyse spectrographique des deux séquences. Pour faciliter la comparaison visuelle des sons à l'octave, un zoom de l'échelle des fréquences des trois sons sol_2 a été appliqué afin que l'étagement des composantes spectrales se présente de la même façon sur les deux analyses. La grille en pointillé figure les harmoniques de 200 Hz pour le sol_2 et ceux de 400 Hz pour le sol_3[31]

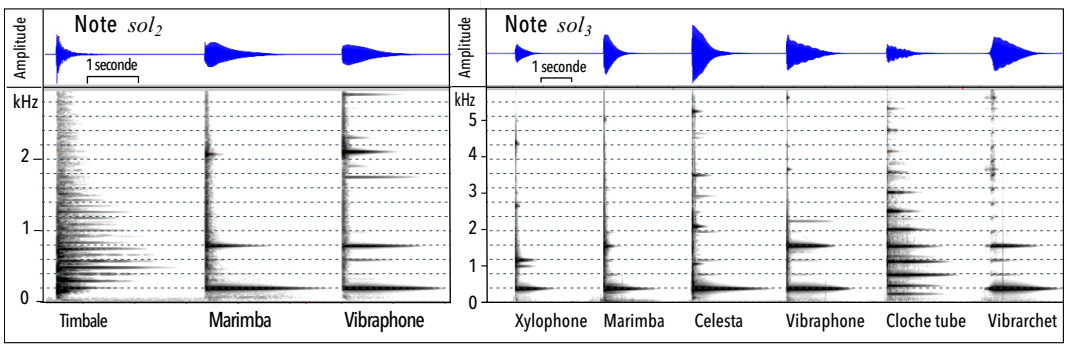


Figure 7.8 Analyses comparatives de sons d'instruments à percussion de hauteur déterminée. À gauche, note sol_2 (échelle de fréquences 0-3 kHz) ; à droite, note sol_3 (échelle de fréquences 0-6 kHz). Exceptées la timbale et la cloche tube, tous les instruments ont une composante intense à la fréquence fondamentale. Les lignes en pointillé sont alignées sur les harmoniques des notes sol_2 et sol_3.

La timbale et la cloche tube ont des formes spectrales caractéristiques sur les sonagrammes, car la répartition spectrale des modes propres d'une membrane tendue et d'un tube rigide sont très différentes de celle des barres vibrantes communes aux autres instruments. Nous pouvons déjà en conclure qu'une des caractéristiques fortes du timbre de ces deux instruments, indépendamment de leur allure temporelle, tient à leur composition spectrale, et plus particulièrement au fait que la perception de hauteur nécessite une écoute spécifique. Les autres sons, tous produits par une lame de bois ou de métal, ont pour caractéristique d'avoir une composante d'intensité notable à l'emplacement de la fréquence fondamentale (1re ligne en pointillé). C'est cette composante qui donne la note (voir chapitre 6, § 3.3).

30. Le « vibrarchet », lame de vibraphone frottée par un archet, n'est pas à proprement parler un son de percussion mais fait partie des sons exploitables dans cet instrumentarium ; il donne l'occasion de vérifier le fait que le spectre d'un instrument entretenu est dépendant des modes propres potentiels qui sont ceux du vibraphone percuté.
31. Pour plus de simplicité, nous assimilons sol_2 à 200 Hz et sol_3 à 400 Hz.

La note d'un instrument à lame vibrante est donc donnée par la fréquence du premier mode vibratoire P1, ce qui en pratique impose un point de frappe particulier. Les fréquences des autres modes, plus notables dans les métaux que dans les bois, contribuent à la caractérisation du timbre causal mais, du fait de l'inharmonicité, peuvent aussi perturber l'écoute des hauteurs. La fréquence du second partiel d'une lame vibrante est comprise entre trois et quatre fois celle de P1. C'est une note « fausse » qui, pour un son de 500 Hz (environ do_4), se situe entre 1 500 et 2 000 Hz (environ la_6 ou sib_7). Les facteurs l'accordent au plus près de la fréquence d'un harmonique de P1[32] : dans notre exemple, c'est $3 \times$ P1 pour le xylophone, $4 \times$ P1 pour le marimba et le vibraphone.

Son 7.10 (20")

Octaves

Lorsque ce partiel est dans la zone d'écoute (voir Glossaire), en couleur sur la figure 7.9, il s'entend comme note autonome et peut, dans certains contextes mélodiques, entrer en concurrence avec le partiel 1. L'exemple sonore **7.10** joue sur de telles ambiguïtés de hauteur dans une suite d'octaves et d'unissons.

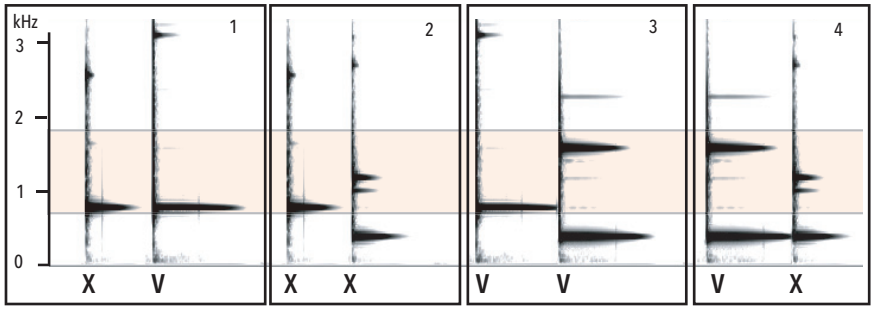


Figure 7.9 Analyse spectrographique d'une séquence jouée alternativement sur un xylophone (x) et sur un vibraphone (v). Paires 1 et 4 : unissons des deux instruments sur sol_4 et sol_3. Paires 2 et 3 : octaves descendantes sol_4-sol_3 sur le même instrument.

En progressant dans l'aigu de la tessiture, l'effet des partiels supérieurs du xylophone[33] et du glockenspiel s'atténue rapidement, car le fondamental masque les partiels plus élevés qui s'estompent perceptivement. Par ailleurs, l'adjonction de tubes résonateurs accordés sur la fréquence du partiel 1 renforce aussi l'intensité du fondamental tout en prolongeant la durée du son.

b/ Les cordes libres frappées et pincées

Du point de vue de la typologie acoustique, la corde frappée ou pincée, dont le son est dû à une impulsion, est une percussion. Du point de vue musical, c'est d'abord une « note ». Les instruments à cordes de la musique occidentale – construits pour jouer des musiques mélodiques et harmoniques – sont travaillés par les luthiers et les musiciens pour tendre vers la quasi-harmonicité des modes et favoriser une bonne longueur de son, et pour atténuer le début percussif. La durée d'extinction dépend principalement du coefficient d'amortissement du matériau dont est fait la corde (métal, fibre végétale, soie, coton, Nylon, Téflon) mais aussi du couplage de la corde avec la structure sur laquelle elle est nécessairement tendue. En vibrant, l'instrument rayonne un son dont le spectre est constitué des fréquences des modes propres de la corde, elles-mêmes transformées en amplitude par celles des composantes de la structure (table, caisse, etc.).

32. Voir Chaigne & Kergomard, 2008, chap. 3, § 4.1.3.
33. Le premier xylophone introduit à l'orchestre ne commençait qu'au do_4 (500 Hz), zone à partir de laquelle la « note » de la lame n'est plus ambiguë.

Au moment de l'impact, on trouve aussi les fréquences propres de la caisse et, lorsque celle-ci est en partie fermée, celles de la cavité dite de Helmholtz (voir Glossaire). À la différence des lames de bois ou de métal, le choix du point d'excitation d'une corde permet de modifier le contenu spectral d'une note sans changer de façon notable la hauteur perçue. Cette nouvelle possibilité ainsi que le fait que cette famille d'instruments est fondamentalement orientée sur la production de notes précises nous conduit, du double point de vue perceptif et musical, à rapprocher les cordes pincées et frappées de la catégorie des instruments à sons entretenus.

Son 7.11 (20″)

Note la_3*, comparaison de la percussion de trois lames puis de trois cordes*

d/ Comparaisons des instruments à lames et à cordes

La figure 7.10 rassemble l'analyse spectrale de six instruments de hauteur déterminée produisant la même note : la_3 (Son **7.11**). À gauche, cette note est jouée sur trois instruments à lame percutée et à clavier (vibraphone, marimba, célesta). À droite, on peut voir les analyses du même la_3 joué sur un piano, un clavecin et une guitare.

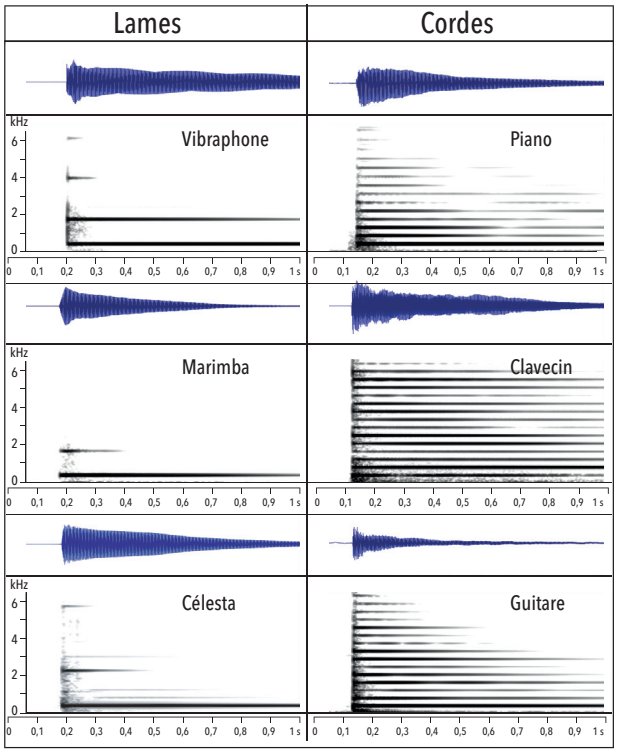


Figure 7.10 Analyse de la note la_3 permettant de comparer la composition spectrale de sons produits par la percussion de barres (partie gauche de la figure) et celles résultant de l'excitation d'une corde frappée ou pincée (partie droite). Durée : 1 s. Analyse à 7 kHz, 1 024 pts.

L'opposition entre ces deux mondes sonores est frappante. Le spectre d'une note produite par une lame vibrante se réduit à deux composantes, dont la plus grave donne la note proprement dite, alors que les sons produits par les cordes libres comportent un nombre important de partiels quasi harmoniques.

2.2.3. Bilan de l'étude des percussions

Nous venons d'examiner un groupe d'instruments dont le son est produit de façon impulsionnelle : aussitôt émis, le son décroît, la structure excitée continuant à vibrer sur les fréquences de ses modes propres. Le classement des musiciens correspond à la typologie acoustique que nous avons présentée, qui est fondée essentiellement sur les propriétés mécaniques des structures excitées : forme géométrique, caractéristiques des matériaux, dimensions globales. Ces objets sont reconnaissables moins au transitoire qui leur est commun, qu'à la forme spectrotemporelle qui combine l'enveloppe dynamique de la décroissance, la durée et la signature spectrale.

La figure 7.11 présente une typologie acoustique schématique des principales familles de sons de percussion.

Les sons du groupe de hauteur indéterminée (partie gauche de la figure 7.11) possèdent chacun une forme spectrotemporelle très caractéristique sur une analyse spectrale, ce qui traduit le fait qu'ils sont immédiatement reconnaissables à l'oreille. Selon Pierre Boulez, « un tam tam existe en lui-même et possède une force centrifuge très puissante » (1984, entretien Claude Samuel), ce qui pose des problèmes d'écriture en particulier pour l'intégration à un orchestre. En revanche, de tels sons bien individualisés sont recherchés en polyrythmie (voir l'exemple de la musique vietnamienne analysée au chapitre 4, § 3.8.4).

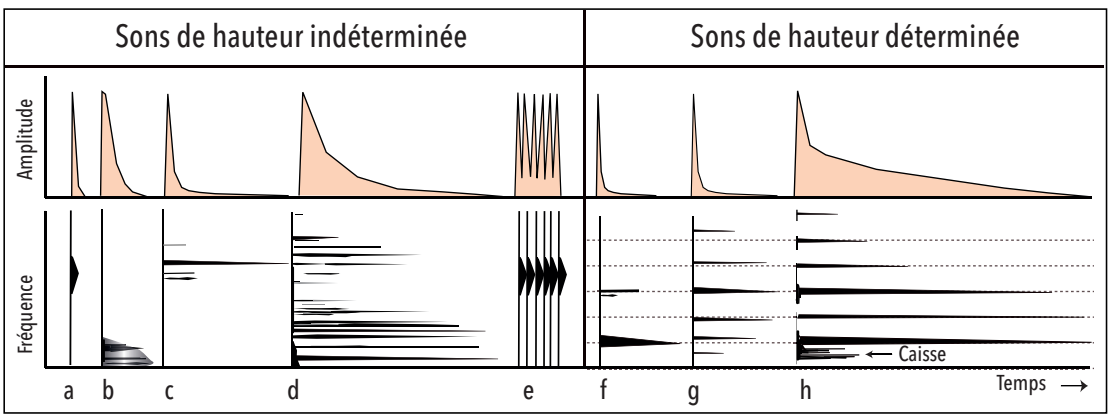


Figure 7.11 Typologie schématique de la forme spectrotemporelle de sons de percussion. À gauche, le groupe des sons de hauteur indéterminée qui diffèrent par la durée d'extinction, la composition spectrale et la tessiture : a, clave ; b, grosse caisse ; c, triangle ; d, cymbale ; e, itérations sur clave. À droite, le groupe des sons d'instruments accordables : f, hauteur due au partiel 1 (barre) ; g, hauteur de type « modèle harmonique » (cloche tube) ; h, spectre quasi harmonique (corde percutée ou pincée) avec réponse spectrale de la caisse.

Le groupe des sons de percussion de hauteur déterminée (partie droite de la figure 7.11) marque une frontière forte, aussi bien instrumentale que perceptive, en raison de la production d'une note. Qu'il s'agisse d'une lame percutée (f), d'un tube frappé (g) ou d'une corde (h), ces instruments servent une musique mélodique et polyphonique. L'écoute des auditeurs s'oriente plus sur les « notes » émises que sur les caractéristiques identitaires de l'objet mis en vibration. Il y a lieu toutefois de faire une place à part aux cordes vibrantes.

2.2.4. Du timbre de l'instrument au timbre de la note

La hauteur produite par une corde étant beaucoup plus robuste que celle produite par un tube, une lame ou une membrane, le musicien peut varier les qualités de timbre d'une note donnée en diversifiant les modes de jeu.

Voici, à titre d'exemple (figure 7.12, Son **7.12**), un mi_3 de guitare (330 Hz) joué de six façons différentes par le même musicien, sur le même instrument, dans la même nuance *mf*. Pour les trois premières notes, jouées sur la même corde, le musicien a modifié le point de pincement ; pour les trois suivantes il a choisi une autre corde de la guitare et changé le mode de jeu (jeu en harmonique effleuré).

Son 7.12 (56")

Guitare : plusieurs modes de jeu pour la même note

Mis à part les changements notables des sons 3 et 6, ces variations peuvent sembler subtiles aux oreilles de certains alors qu'elles sont capitales pour des auditeurs experts. Comparez les attaques, le nombre de partiels, la décroissance du son et les « bruits » de caisse. Le zoom fréquentiel sur la partie basse du spectre met en évidence les changements de bruits d'attaque selon la corde et selon le point de pincement.

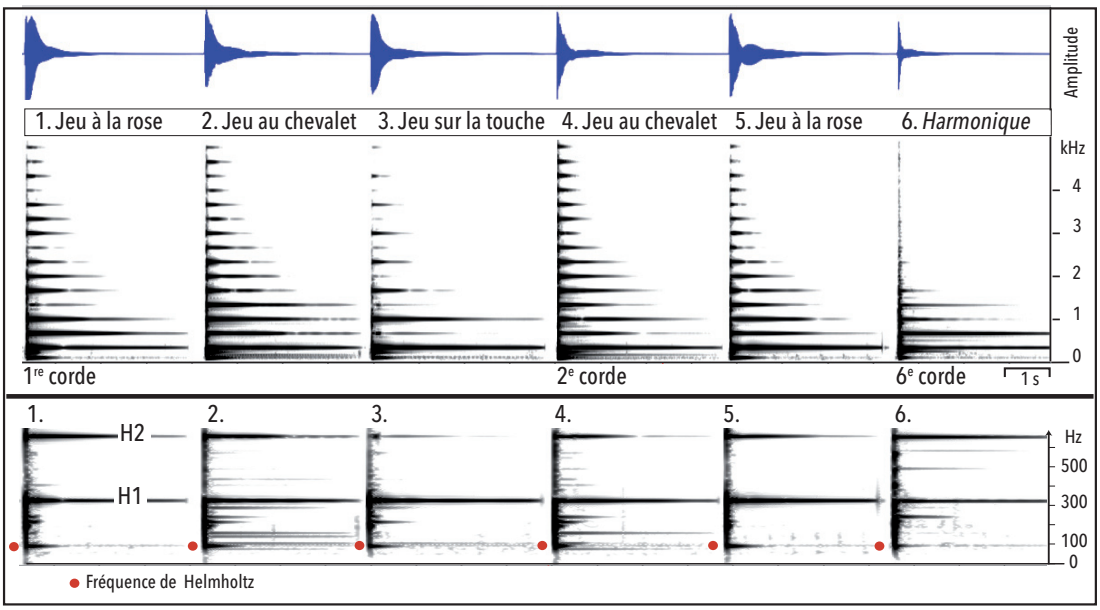


Figure 7.12 Analyse d'un mi_3 joué de six façons différentes. En haut : enveloppe dynamique et sonagramme. (FFT : 2 048 pts). Remarquez particulièrement le spectre de composantes impaires du son 3 : jeu sur la touche et dominance de l'harmonique 2 dans le jeu en « harmonique effleuré ». En bas : zoom fréquentiel sur la partie basse du spectre. On distingue, sous la fréquence du fondamental (H1), plusieurs composantes des modes de caisse excitées par l'attaque, et en particulier la fréquence de résonance dite de Helmholtz, à 115 Hz.

En conclusion, la note d'un instrument de hauteur définie peut donc changer de « timbre » de façon assez considérable, tout en restant inscrite dans le « timbre » de l'instrument. Dans le groupe des instruments de hauteur définie, il existe donc bien deux niveaux de constitution du timbre : celui qui relève du principe de production mécanique du son (ici, corde fixée sur une caisse et pincée au doigt) et celui des variations de production propres à cet instrument qui, sous le contrôle de l'interprète, sont l'essence même du jeu musical.

À l'image des poupées russes, les variations de timbre des notes sont encapsulées dans la forme acoustique du timbre identitaire de l'instrument, celui dont la permanence assure une stabilité à l'écoute et qu'il s'agit maintenant de caractériser.

2.3. Les sources sonores de hauteur définie

L'étude doit être envisagée selon différents points de vue.

1/ Du point de vue organologique, la production d'une note résulte de l'interaction entre un geste fournisseur d'énergie et une structure oscillante et, comme pour les percussions, ce sont les événements sonores de l'enveloppe spectrotemporelle qui caractérisent le timbre causal. En nous fondant sur le fonctionnement acoustique des sources sonores mécaniques, nous pourrons définir le cadre d'une typologie générale.

2/ Du point de vue perceptif, le fait que les hauteurs (les notes) captent prioritairement l'attention des auditeurs, complique l'étude des aspects qualitatifs du timbre, car des relations complexes existent entre hauteur, contenu spectral et tessiture.

3/ Du point de vue musical, les instruments sont conçus pour « jouer », pour varier les sons afin de les rendre vivants et intéressants à l'écoute. Pour la plupart d'entre eux, particulièrement les instruments mélodiques, c'est l'interprète qui « fait » le son. Il nous faudra donc explorer, selon l'expression d'Émile Leipp, le « champ de liberté » de ces variations, faire la part de ce qui vient de l'instrument et de ce qui dépend du jeu de l'instrumentiste, car certains modes de jeu, comme le vibrato, peuvent participer du timbre causal alors que d'autres ne relèvent que du goût et du style de jeu du musicien.

Notre objectif sera de saisir les indices acoustiques les plus saillants et les plus robustes, ceux qui permettent la catégorisation en temps réel d'une seule note, nous réservant d'aborder ensuite les variations dues au changement de tessiture. Finalement nous pourrons approcher la notion d'instrument, notion inséparable de la musique qu'il sert. Par ce terme, nous désignons non seulement l'objet dans son principe de fonctionnement acoustique, mais aussi la famille des sons utilisables et adaptés à une musique (voir § 2.5.5 et 2.5.4). Des expériences d'écoute nous permettront d'évaluer, dans chaque domaine, ce qui permet de caractériser les divers aspects du timbre en vue d'interpréter les analyses acoustiques.

2.3.1. Une typologie acoustique élémentaire des sources de hauteur définie

Prenons pour point de départ le fonctionnement des sources sonores.

Pour produire des sons de hauteur bien définie – périodique ou quasi périodique –, nous disposons de deux modes de fourniture d'énergie qui sont l'impulsion (frappe, pincement) et l'entretien (frottement, souffle), et de trois grands systèmes producteurs de « notes » dans lesquels un système oscillant est couplé à une structure (figure 7.13). Il s'agit, d'une part, du système de la *corde* associée à une caisse (tous les instruments dits à corde) ou à une cavité (arc musical), et, d'autre part, des deux systèmes aériens du *jet d'air* (flûtes) et de l'*anche* (simple, battante, libre, lippales[34]), couplés à diverses sortes de tuyaux (cylindriques, coniques, percés ou non de trous latéraux).

34. Terme employé pour désigner les lèvres, c'est-à-dire des anches élastiques, susceptibles de tension comme le sont les mal nommées « cordes » vocales.

Selon les instruments, la structure contrôle la fréquence d'oscillation à des degrés divers, en assure la stabilisation et contribue au rayonnement du son. La combinaison de ces éléments donne lieu à une grande variété de réalisations instrumentales. Pour envisager une typologie pertinente du point de vue perceptif, il importe de poser la question du nombre d'instruments différents qui sont identifiables à l'écoute d'un seul son, sans le secours de la vue ni même du style de musique. Nous savons que dès les premières millisecondes s'opère la séparation fondamentale entre les sons produits par impulsion et les sons entretenus (voir chapitre 4). Les premiers ont une durée limitée : sitôt émis, ils ne font que décroître en amplitude jusqu'à extinction alors que pour les sons entretenus, l'amplitude sonore et la durée varient au gré de l'instrumentiste. Cette dichotomie, qui correspond aussi pour partie à une frontière entre instruments à cordes polyphoniques et monodiques, se retrouve au premier niveau des résultats de tests d'écoute qui mélangent les sons de percussion et les sons entretenus (Iverson *and coll.*, 1993). Deux groupes de sources forment la catégorie des sons entretenus : les cordes et les vents, eux-mêmes divisés en flûtes et anches. Ces distinctions sont-elles pertinentes du point de vue perceptif ? Sans répondre immédiatement à la question, l'exemple sonore suivant pose le problème.

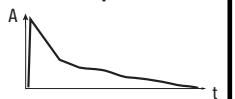
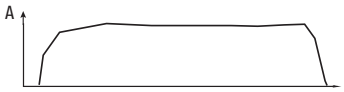
		I - Impulsion	II - Oscillations forcées : sons entretenus	
1	**Enveloppe Dynamique**	 *Amplitude décroissante*	 *Amplitude variable*	
2	**Fournisseur d'énergie**	Frappe, Pincement *Choc*	Archet *Frottement*	Débit aérien *Souffle*
3	**Producteur de note** (oscillateur & structure couplée)	Corde Tendue sur une caisse	Jet d'air oscillant ou Anche(s) Associés à tuyaux ou cavités	

Figure 7.13 Typologie acoustique de base de la production de sons de hauteur définie.

2.3.2. Deux expériences sur la reconnaissance de sons entretenus isolés

Son 7.13 (8")

Son répété trois fois

L'exemple sonore 7.13 donne à entendre un son entretenu répété trois fois.
Par quel instrument a-t-il été produit ?

IMPORTANT : Il faut écouter l'exemple sonore avant de lire la suite.

Un son intriguant

À l'écoute de cet exemple, les auditeurs les plus divers, y compris des musiciens chevronnés, citent des instruments comme la clarinette, le saxophone et, plus fréquemment, l'harmonica ou une trompe d'auto. Tous sont des instruments à vent de la famille des anches. Or, les deux instruments auxquels nous avons emprunté séparément le transitoire d'attaque et la partie stable pour réaliser cette chimère ne sont jamais mentionnés. Pour comprendre ce phénomène surprenant (dont nous donnons la clé au § 2.3.6) il faut examiner plus en détail les événements caractéristiques de l'identité sonore des différentes sources instrumentales : c'est l'objet du test d'écoute que nous avons mis en place.

Test de reconnaissance de sons instrumentaux isolés

Dans un premier temps, nous allons comparer l'analyse acoustique de sons issus d'instruments différents jouant la même note. Plusieurs bases de données offrent des exemplaires de sons isolés pour les principaux instruments de l'orchestre. Toutefois, lorsqu'on écoute dans un ordre quelconque les sons prélevés dans une base sans connaître le nom du fichier, l'expérience aboutit toujours au même résultat : parmi la quinzaine de sons retenus, il y en a toujours deux ou trois dont l'identification est problématique, voire impossible. Avant d'entreprendre une analyse acoustique de sons instrumentaux, il est donc nécessaire de s'assurer que ceux que nous choisissons sont bien représentatifs de chaque instrument, en effectuant un test de reconnaissance avec des auditeurs.

Nous avons puisé les sons de notre expérience dans quatre bases de données disponibles[35] et, pour rester au plus près du signal acoustique, nous n'avons retenu que des sons tenus joués sans vibrato. Notre corpus de sons instrumentaux est constitué de sons donnant la même note : 60 sons do_3 (261 Hz) et 60 sons la_3 (440 Hz) qui diffèrent par la durée, l'intensité, le modèle d'instrument, le joueur et les conditions d'enregistrement. Huit auditeurs (cinq musiciens expérimentés et trois musiciens amateurs), informés qu'il s'agissait de sons d'instruments de musique extraits de bases de données, ont aimablement accepté de se prêter au test. Il leur était demandé de reconnaître, « dès la première écoute », l'instrument ou la famille, indépendamment des disparités manifestes (en particulier le diapason et l'intensité). Ils n'avaient pas connaissance de la liste des instruments que voici : violon, alto, violoncelle ; flûte traversière en *ut*, flûte à bec, flûte de Pan ; hautbois, cor anglais, basson ; clarinette *sib* ; saxophone alto, saxophone ténor ; trompette, trombone ténor, cor, tuba, accordéon. Les résultats de ce test nous ont fourni une première taxinomie des sources de hauteur définie fondée en perception.

2.3.3. Les résultats du test d'écoute de sons isolés

1/ Quelles que soient les conditions d'écoute (casque ou haut-parleurs) ou l'expertise musicale des sujets, les sons générés par les trois types d'oscillateurs « corde, jet d'air, anche » ont été parfaitement reconnus et ne se confondent pas entre eux (figure 7.14, niveau 3).

2/ À l'intérieur de la catégorie « anche », deux groupes ont été identifiés par tous les musiciens. Les « cuivres » (anches lippales), au sein duquel les sons de trompette sont identifiés pour les deux hauteurs (do_3 et la_3), et les « sons d'anche libre » (accordéon) qui forment également une catégorie ne se mélangeant pas aux autres. Pour l'ensemble des auditeurs, il existe donc quatre catégories non équivoques de sons entretenus : cordes frottées, flûtes, anches lippales et anches libres (figure 7.14 niveau 4) ; la cinquième catégorie (anches roseau) regroupe les instruments restants.

3/ La surprise concerne cette cinquième catégorie, celle des anches en roseau ou encore à « trous latéraux », qui rassemble hautbois, clarinette et saxophone, que seuls les musiciens entraînés ont pu discriminer, avec tout de même quelques erreurs et bien plus difficilement pour le la_3 que pour le do_3. Certains sont parvenus également à identifier les sons de basson et de cor alors que les sons d'instruments d'une même famille (alto et violoncelle pour les cordes et trombone et tuba pour les cuivres) ont presque toujours été confondus (figure 7.14, niveau 5).

35. Base McGill (Opolko,1989) ; base RWC (Goto, 2003) ; base Iowa (Fritts, 1997) ; base SOL (© Ircam, 2009). Nous adopterons respectivement les sigles suivants pour les désigner : McG, RWC, Iow, Orch.

Ces résultats recoupent en partie ceux des chercheurs ayant effectué des tests de reconnaissance de sons isolés (Berger, 1964 ; Saldanha & Corso, 1964 ; Clark, Robertson & Luce, 1964 ; Martin, 1999) ainsi que ceux obtenus avec les algorithmes de traitement de signal mis en œuvre dans les programmes MIR[36]. Il est cependant difficile d'effectuer des comparaisons chiffrées avec ces travaux, car le nombre et le type d'instruments retenus change d'un auteur à l'autre et surtout parce que les taux de reconnaissance et les confusions sont généralement comptabilisés pour chaque instrument individuellement, sans prendre en compte la proximité des instruments dans certaines familles (violon et alto ou hautbois et cor anglais, par exemple).

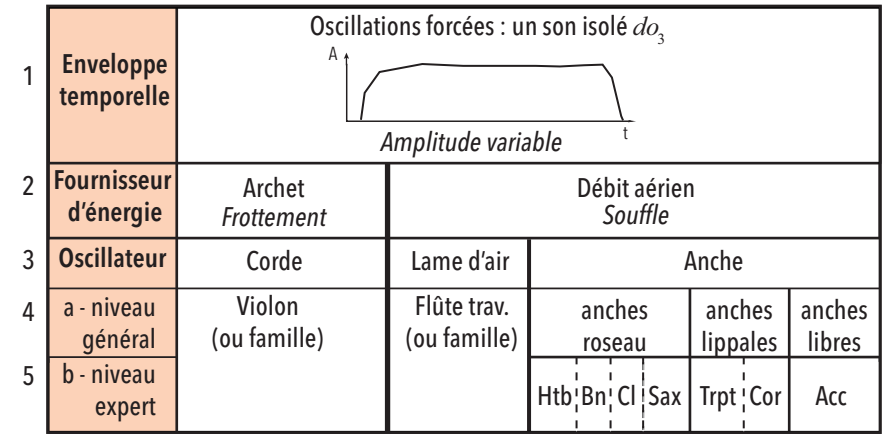


Figure 7.14 Taxinomie hiérarchique des sources instrumentales rendant compte des résultats du test de reconnaissance portant sur deux séries (do_3 et la_3) de chacune 60 sons isolés issus des bases de données. Codes des noms d'instruments : Htb (hautbois), Bn (basson), Cl (clarinette), Trpt (trompette), Cor (cor), Acc (accordéon)

2.3.4. Le transitoire d'attaque et la formation de la note : une analyse temporospectrale

À l'issue de ce test, nous avons sélectionné, pour chaque note do_3 et la_3, deux sons ayant obtenu le meilleur score de reconnaissance par instrument, que nous pouvons donc considérer comme de bons représentants de chaque catégorie. Compte tenu de la tessiture choisie, nous n'examinerons que les sept instruments jouant le soprano : violon, flûte traversière, hautbois, clarinette, saxophone, accordéon et trompette (Son **7.14**). Les sons analysés étant plus représentatifs d'un mode de production sonore donné – corde frottée, embouchure de flûte, etc. – que d'un instrument de musique proprement dit, nous emploierons souvent le numéro d'ordre de la figure.

Son 7.14 (58")

Le son do_3 produit par sept sources instrumentales différentes (2 séries). L'ordre d'écoute correspond à une lecture horizontale de la figure (un son par catégorie).

La figure 7.15 est une analyse temporospectrale des 300 premières millisecondes de chaque son do_3. Malgré la disparité des techniques de jeu, des modèles d'instrument et des conditions d'enregistrement, il est possible d'apprécier au premier coup d'œil les allures communes aux sons situés l'un au dessous de l'autre, produits par le même type de source

La même analyse réalisée avec les sons du la_3 (Son **7.15**) permet en outre d'apprécier la stabilité des caractéristiques sur une note plus élevée (figure 7.16).

Son 7.15 (58")

Le son la_3 produit par sept sources instrumentales différentes (2 séries). L'ordre d'écoute correspond à une lecture verticale de la figure (deux sons par catégorie).

36. MIR pour *Music Information Retrieval*. Voir figure 7.41.

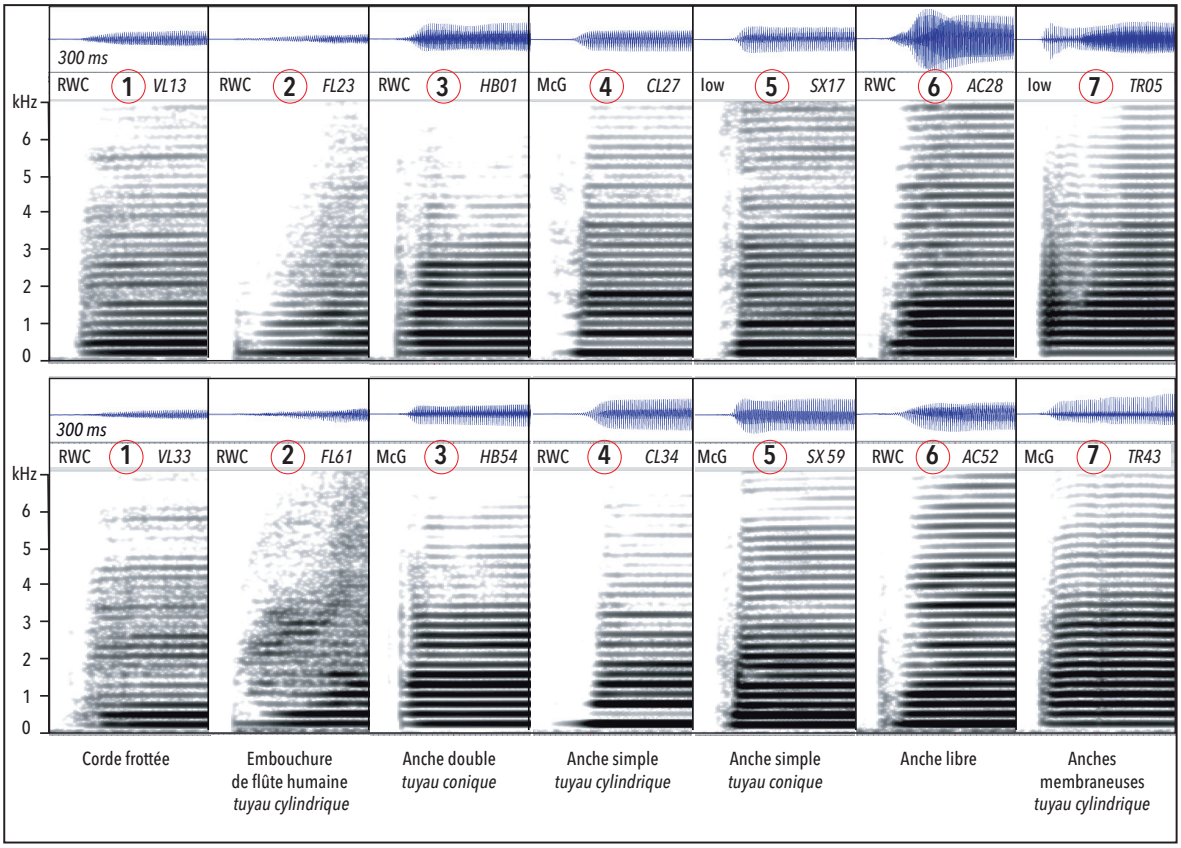


Figure 7.15 Note do_3. Analyse de la partie initiale (300 ms) des paires de sons représentant les sept catégories de sons instrumentaux identifiées par les musiciens (analyses : 7 kHz ; 1 024 pts).

Son 7.14 (58")

Un son par instrument (2 séries)

En comparant les événements initiaux sur les deux figures 7.15 et 7.16, on remarque des spécificités notoires.

- Sons 1 des cordes frottées : le transitoire s'établit d'emblée sur tout le spectre en mixant bruit d'archet et harmoniques.
- Les sons 7 de trompette ont en commun de concentrer l'énergie dès le début sur le groupe des premiers harmoniques dont l'établissement s'effectue avec de petites oscillations en fréquence (voir figure 7.17, page 314), les harmoniques supérieurs n'apparaissant qu'en second lieu, avec un décalage temporel caractéristique.
- Au centre, les sons 2 à 6 présentent des bruits de souffle variés selon le type de coup de langue. Ces bruits connectés au démarrage du fondamental (surtout les sons 4, 5, 6) précèdent nettement l'arrivée des harmoniques supérieurs.
- Les sons 2 du do_3 de flûte traversière ont l'allure caractéristique de la note la plus grave de l'instrument, dont on sait que la mise en route est souvent problématique. Ceux du la_3 montrent le décalage temporel entre le bruit de souffle de l'attaque et l'établissement des harmoniques.

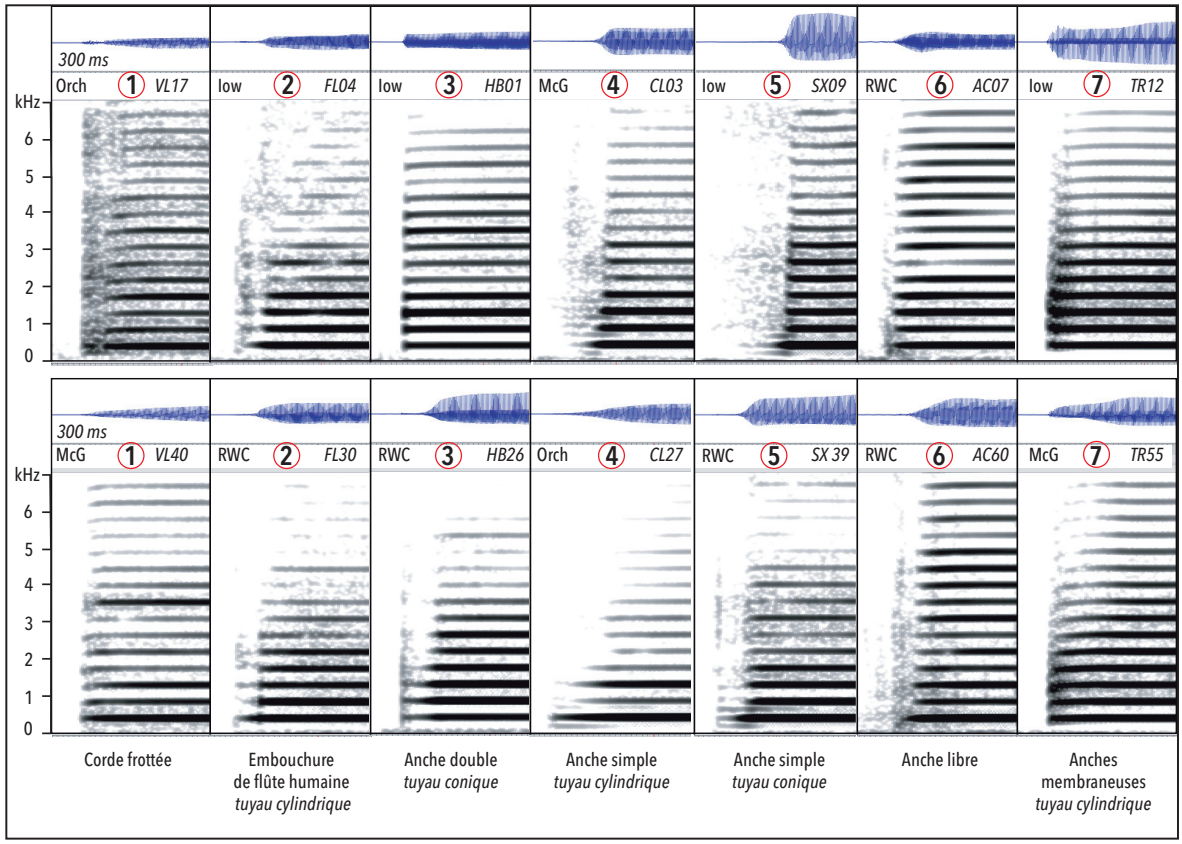


Figure 7.16 Note la_3. Analyse de la partie initiale (300 ms) des paires de sons représentant les sept catégories de sons instrumentaux isolés identifiées par les musiciens (analyses : 7 kHz ; 1 024 pts)

Compte tenu du fait que les sons analysés sont tous reconnus par les musiciens, nous pouvons en conclure que leur « forme spectrale », c'est-à-dire les bruits, leur décalage avec l'arrivée des harmoniques supérieurs et les instabilités de fréquence est un critère plus pertinent que la durée globale du transitoire qui change d'un interprète à l'autre, comme on peut l'observer.

Son 7.15 (58")

Deux sons par instrument (une série)

Cette durée varie aussi beaucoup selon l'intensité (voir chapitre 5, § 5) et selon l'expressivité du jeu.

Toutefois, pour certains instruments comme la flûte et l'accordéon, la durée d'établissement du son ne peut descendre en dessous d'une valeur minimale, ce qui constitue en soi un critère de discrimination. Le hautbois (son 3), réputé pour sa précision d'attaque, produit la note de façon quasi instantanée sur l'ensemble du spectre (note la_3, son supérieur).

Les caractéristiques remarquables de l'établissement des sons de hauteur définie

Son 7.16 (7")

Deux sons isolés : trompette, flûte traversière

1/ Le premier niveau de catégorisation se fonde sur des indices globaux, temporels et énergétiques ; c'est la discrimination entre souffle ou archet combinée à celle d'un spectre large ou réduit qui oriente vers les trois groupes : cordes, flûtes, anches.

2/ Le second niveau concerne la façon dont s'établit la note : a) rapidement et précisément (corde, hautbois) ; b) rapidement avec oscillations (cuivres) ; c) progressivement avec des fréquences étrangères (flûtes et tuyaux d'orgue) ; d) progressivement et sans fluctuation (anches battantes, anche libre). Le zoom temporel effectué sur les analyses de deux sons de trompette et de flûte (Son **7.16**) met en évidence deux de ces caractéristiques (figure 7.17). Les changements spectraux présents dans le transitoire sont au moins aussi importants que la seule durée, généralement retenue dans les études.

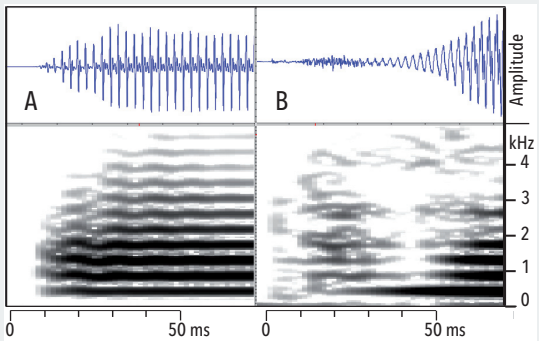


Figure 7.17 Zoom temporel sur deux transitoires typiques (note la_3). A : anches lippales (trompette, son Iowa TR12) ; B : jet d'air oscillant (flûte traversière, son Iowa Fl04). Échelles : 5 kHz, durée à l'écran, 70 ms. FFT : 512 pts.

3/ Les caractéristiques spectrales proprement dites, c'est-à-dire les rapports d'énergie entre les harmoniques de la note sur la base desquels se différencient les instruments du groupe « anches de roseau », nécessitent une écoute harmonique donc une éducation musicale plus avertie dont ont fait preuve les musiciens expérimentés de notre test.

2.3.5. Le contenu harmonique de la note : les analyses spectrales moyennées (LTAS)

Pour chacun des 28 sons représentant les sept catégories sonores les plus reconnaissables, nous avons effectué un spectre moyenné sur toute la durée du son[37]. L'analyse est limitée à 6 kHz, ce qui correspond à 22 harmoniques pour le do_3 et à 13 harmoniques pour le la_3. Les quatre catégories les plus facilement identifiées sont regroupées sur la figure 7.18. Les spectres présentés prennent en compte le fonctionnement global de la source : bruit d'écoulement ou de frottement, résonances de caisse et modes propres des tuyaux, particularités du rayonnement, variabilité des prises de son et des modèles d'instruments utilisés. Malgré les disparités visibles pour les sons d'une même colonne, des tendances se dégagent lorsqu'on compare le nombre total d'harmoniques, l'enveloppe des spectres et la distribution d'énergie sur les premiers harmoniques (voir chapitre 2, § 4.1 pour l'écoute de trois spectres types).

37. Spectre LTAS, Long-term average spectrum, effectué avec Wavesurfer, FFT 4 096 pts (voir chapitre 2, § 3.5.2).

Enveloppe spectrale globale

À première vue, toutes les sources des figures 7.18 et 7.19 produisent un spectre riche en harmoniques de 0 à 6 kHz, excepté la flûte traversière pour laquelle l'amplitude chute rapidement dans le bruit de fond au-delà du 6e ou 8e harmonique. Cette opposition rend compte de la position singulière des sons de flûte dans la catégorisation instrumentale.

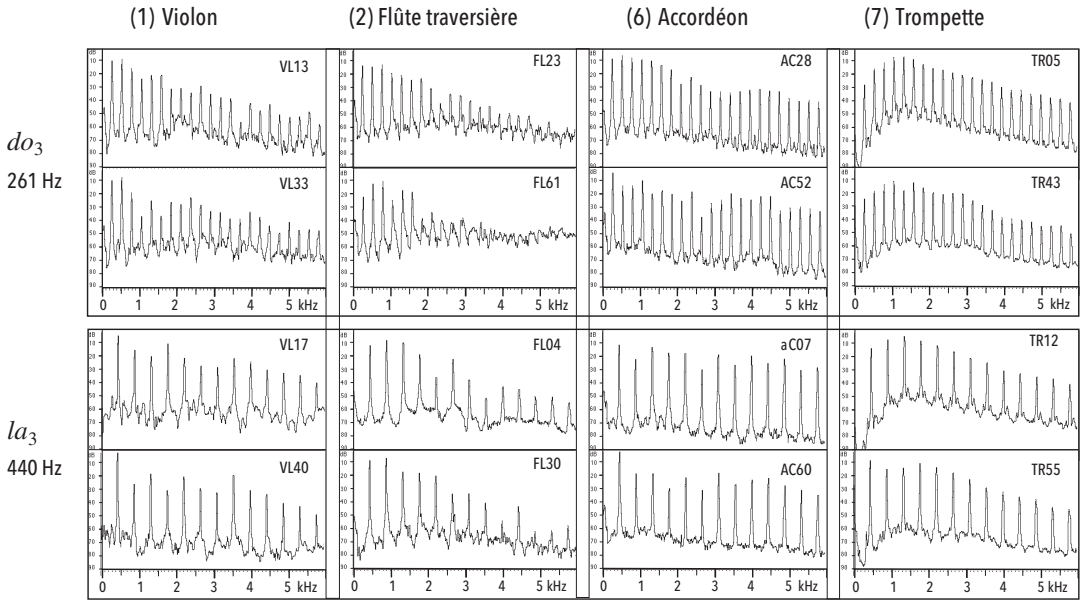


Figure 7.18 Spectres moyennés (LTAS) sur la durée de chaque son de l'expérience d'écoute. Analyse des quatre catégories de sources aisément reconnues par tous les auditeurs du test d'écoute (voir § 2.3.3). Échelle verticale : 90 dB. Échelle horizontale : 6 kHz.

On peut faire une distinction supplémentaire au sein des sons riches en harmoniques en observant le taux de décroissance de l'intensité entre le grave et l'aigu, plus particulièrement pour la note do_3. Ces considérations valent pour des spectres homogènes dans lesquels les composantes sont alignées de façon régulière.

Dans la réalité, l'enveloppe spectrale est affectée de bosses (ou formants) qui peuvent englober plusieurs harmoniques ou montrer d'importantes irrégularités d'intensité d'un harmonique à l'autre. Nous sommes particulièrement sensibles à ces dernières lorsqu'il s'agit des premiers harmoniques, de 1 à 6. En effet, ces harmoniques étant séparés par un intervalle supérieur au tiers d'octave « tombent » dans des bandes critiques différentes (voir chapitre 3, § 2.2) et sont donc perceptibles individuellement.

Examinons en tout premier lieu le fondamental (H1) puis les intensités relatives entre harmoniques pairs et impairs.

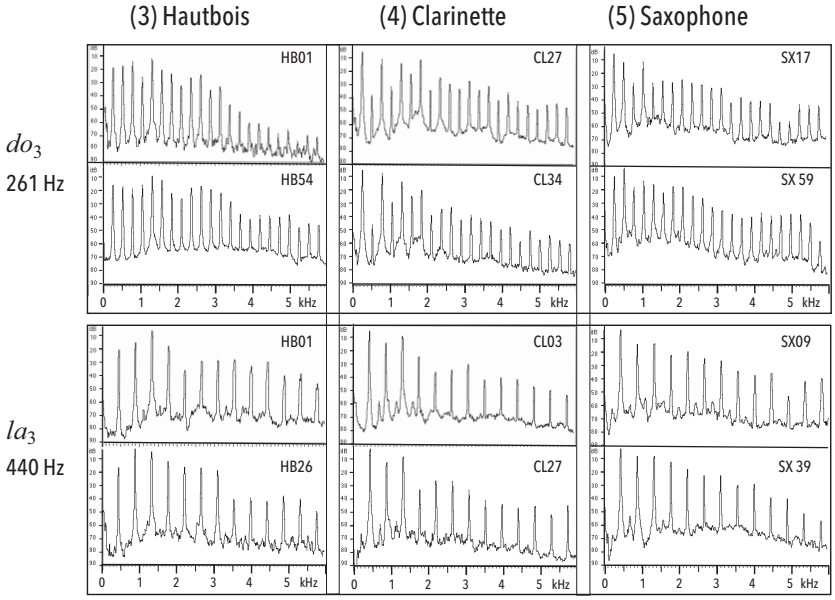


Figure 7.19 Spectres moyennés (LTAS) sur la durée de chaque son de l'expérience d'écoute. Analyse des trois catégories de sources – les instruments à anche à trous latéraux – que les auditeurs du test d'écoute reconnaissent difficilement (voir § 2.3.3). Échelle verticale : 90 dB. Échelle horizontale : 6 kHz.

Importance relative de l'intensité du fondamental (H1) par rapport à l'ensemble du spectre

L'harmonique 1 (fondamental) domine les autres par son intensité pour la plupart des instruments, à l'exception des sons 7 (trompette) et 3 (hautbois) pour lesquels le maximum se situe plutôt sur les harmoniques 3 à 5, ce qui explique en partie les confusions entre ces deux sources observées dans notre test. Remarquons que dans l'orgue les facteurs et les instrumentistes distinguent trois familles de jeux à bouche (flûtes, principaux, gambes) qui sont caractérisées par la balance énergétique entre le fondamental et les premiers harmoniques. De même les instrumentistes à cordes sont sensibles à l'intensité du fondamental puisqu'ils peuvent ajuster la balance entre fondamental et harmoniques supérieurs en réglant la distance du point d'excitation par rapport au chevalet. Dans le jeu des instruments à embouchure, des variations similaires se produisent avec les sourdines (voir figure 7.21, page 319).

On voit que l'appréciation de l'intensité relative entre les premiers harmoniques peut s'interpréter tantôt comme critère d'identification d'un instrument – le hautbois par rapport à la clarinette et au saxophone dont le fondamental est très intense – et tantôt comme variations de qualité du son d'un instrument donné. C'est donc une clé d'identification moins robuste que les caractéristiques temporelles, d'autant qu'elle est sujette aux variations du rayonnement dans l'espace, et il faut, pour la mettre en œuvre, avoir déjà développé une expertise d'écoute assez fine.

Balance entre harmoniques pairs et impairs

Les harmoniques pairs, 2 et 4, reproduisent la note fondamentale à une et deux octaves de distance alors que les harmoniques impairs, 3 et 5, correspondent à deux notes différentes qui sont respectivement la quinte et la tierce majeure du fondamental. La dominance paire ou impaire est caractéristique du spectre de certains instruments à vent, ce qui correspond à deux fonctionnements acoustiques différents. Les instruments dans lesquels la série des modes propres du tuyau est voisine de la série complète des harmoniques (flûtes, anches et tuyau conique comme le

saxophone) sont des instruments *octaviants*. Ceux pour lesquels la série des modes propres du tuyau est voisine des seuls harmoniques impairs (flûtes de Pan, bourdon d'orgue, anche et tuyau cylindrique comme la clarinette et le cromorne) sont des instruments *quintoyants*. Dans notre expérience, cette caractéristique permet de départager les sons produits par le saxophone de ceux de la clarinette, du moins pour les sons les plus graves du premier régime. La distinction est beaucoup moins nette pour le la_3.

Une distinction similaire existe dans les instruments à cordes, en particulier lorsqu'on peut choisir le point d'excitation. Jouer « sur la touche » ou pincer au milieu de la corde (guitare, harpe) favorise grandement les mouvements vibratoires des composantes impaires au détriment des paires (voir le son 3 de la figure 7.12, page 307). C'est aussi le cas du *muselar*, instrument cousin de l'épinette dont les sautereaux pincent les cordes en leur milieu. Pour certains instruments, l'alternance entre sons à spectres pairs et impairs est recherchée : par exemple (Son **7.17**), le jeu de la guimbarde avec souffle (spectre pair), par opposition au jeu sans débit d'air (spectre impair). Enfin, citons l'exemple de certaines flûtes à bec de facture raffinée dont le réglage d'embouchure permet de passer progressivement d'un spectre impair à un spectre comportant tous les harmoniques en forçant légèrement le souffle.

Son 7.17 (17″)

Jeu de la guimbarde avec et sans souffle (voir page 387)

La balance énergétique entre les harmoniques pairs et impairs joue un rôle important dans la catégorisation spectrale des sons, mais sa reconnaissance relève d'une écoute musicale experte, ce qui explique que seuls les musiciens ont identifié la clarinette ou le saxophone à l'audition d'un son isolé. Il est important de remarquer que, de même que pour le fondamental, la prédominance des harmoniques impairs s'interprète tantôt comme un critère causal (flûte bouchée par rapport à flûte ouverte) et tantôt comme une variation qualitative du son de l'instrument qui relève du mode de jeu (guimbarde, harpe).

L'harmonique 5

L'harmonique 5 fait entendre la tierce majeure du fondamental. Lorsqu'il domine les quatre premières composantes par son intensité, il peut jouer un rôle discriminant dans la catégorisation sonore (voir dans notre exemple les deux notes do_3 du hautbois qui proviennent de deux bases de données différentes). À l'orgue, l'harmonique 5 caractérise le jeu de « cornet » et celui de « tierce ». L'effet est d'autant plus marqué que l'amplitude des harmoniques décroît très fortement à partir du sixième.[38]

Les formants

Pour trois des quatre sons de trompette de la figure 7.18 on remarque une « bosse » de l'enveloppe spectrale aux environs de 1300 Hz : c'est un formant. Le hautbois possède deux formants principaux dont le premier, situé également vers 1300 Hz, explique en partie les confusions de cet instrument avec la trompette. Le deuxième formant, voisin de 2500 Hz, contribue à l'émergence du hautbois dans l'orchestre. Lorsque la fréquence centrale d'un formant est fixe, les harmoniques renforcés changent en fonction de la fréquence fondamentale de la note jouée. Le premier formant de la trompette renforce l'harmonique 5 du do_3 et l'harmonique 3 du la_3.

Il peut arriver que les zones formantiques d'un son instrumental se trouvent dans les mêmes rapports de fréquence que ceux d'une voyelle. Cet effet est recherché dans le jeu d'orgue dit de « voix humaine », mais on ne peut l'obtenir que dans une portion restreinte de la tessiture, celle au sein de laquelle la fréquence fondamentale et les zones formantiques sont dans des rapports de fréquence comparables à ceux de la voix humaine (voir encadré page suivante).

38. Remarque : un harmonique peut dominer s'il a une grande amplitude mais aussi lorsqu'il est situé à une frontière du spectre.

Les formants et le système vocalique

Un formant est une bande de fréquence à l'intérieur de laquelle les harmoniques sont renforcés. Il peut être fixe ou variable en fréquence. Deux paramètres caractérisent un formant : sa largeur de bande Δf et son amplitude, liée au « facteur de qualité » de la résonance.

La notion de formant nous est familière, car elle est caractéristique de la voix humaine (reportez-vous au chapitre 9, § 1.4 pour plus de détails). Le son produit par les cordes vocales traverse des cavités (pharynx, cavités buccale, nasale) dont les résonances transforment l'enveloppe du spectre. L'ajustement de ces résonances permet de produire les voyelles. Le sonagramme *a* montre l'analyse spectrale d'une voix chuchotée féminine articulant cinq voyelles. En l'absence des harmoniques du son glottique, les formants sont plus faciles à identifier.

Les deux premiers formants F1 et F2, qui caractérisent bien les voyelles, sont représentés en phonétique sur le diagramme *b*. Les trois voyelles : « a », « i », « ou » sont dites « cardinales », car elles existent dans toutes les langues. Elles correspondent aux formes spectrales les plus antagonistes : écartement maximal des formants pour le « i » et rapprochement dans le bas du spectre pour le « ou ». Le « a » occupe une position centrale. De part et d'autre du « a » se trouvent « è » et « o ». La position en fréquence du deuxième formant est très discriminante pour la reconnaissance des voyelles. Pour la caractérisation des sons instrumentaux, c'est plutôt la position du premier formant qui est à considérer, car sa zone de variation, qui va de 250 à 800 Hz, correspond à une zone importante de la tessiture musicale, pratiquement celle de la clé de sol (voir la figure c).

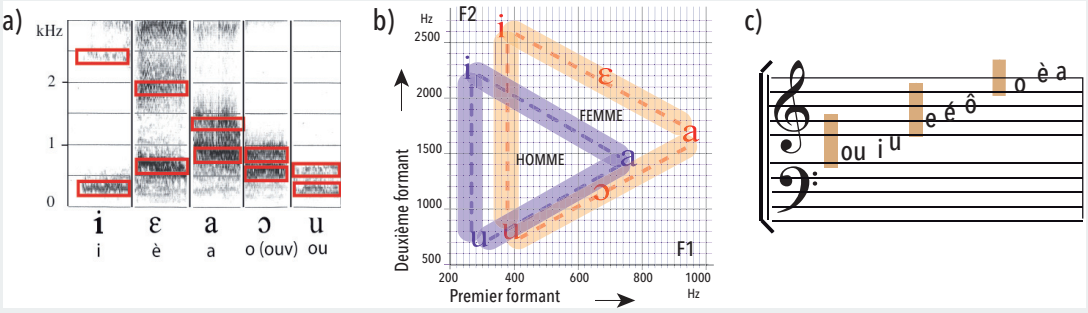


Figure 7.20 Analyses des deux premiers formants caractéristiques des voyelles : « i », « è », « a », « o », « ou ». Analyse a : spectrogramme des voyelles dites en voix chuchotée (voix féminine). Notation IPA (alphabet phonétique international) et notation en français courant. Analyse b : position moyenne des voyelles dans le plan F1-F2 (fréquences des deux premiers formants). En rouge : voix de femme ; en violet : voix d'homme. Analyse c : position approchée du premier formant projeté sur une portée musicale.

Le système vocalique est tout à la fois flou et précis. L'identification d'une voyelle résulte d'une interprétation perceptive, car la voyelle n'a d'existence que lorsqu'elle fait partie d'un système phonologique défini. Les formants vocaliques ne correspondent pas à des valeurs de fréquence fixes. Ils diffèrent d'une langue à l'autre et, pour une même langue, ils diffèrent selon les locuteurs, en particulier entre les hommes et les femmes, figure *b*. Chaque auditeur constitue ses prototypes personnels de référence, tout en acceptant une certaine tolérance pour les écarts, afin de comprendre la langue parlée par des locuteurs différents.

Enfin, nous sommes démunis pour repérer les formants dont les fréquences sortent du système vocalique.

L'approche vocalique du timbre permet par ailleurs d'explorer à l'oreille les zones de résonances mais il ne faut pas en espérer un repérage précis car l'organisation spectrale des voyelles est anamorphosable. Les limites entre catégories vocaliques sont élastiques, chaque type de voyelle se positionnant par rapport à ses collatérales.

Un « anti-formant », ou zone de réjection spectrale, peut également jouer un rôle important dans l'identification des caractéristiques spectrales d'un instrument. Voir sur la figure 7.28, page 328, la gamme chromatique du basson.

Signalons aussi la caractéristique des instruments de la famille du violon dont les zones formantiques sont extrêmement étroites et nombreuses, ce qui produit de nombreuses et fines variations d'une note à l'autre alors que l'allure spectrale d'ensemble est assez régulière. En observant cette caractéristique, Max Mathews a réalisé un violon électronique, joué normalement à l'archet, mais dont la caisse de bois est remplacée par un système de filtres (Mathews, 1977).

Un exemple de transformation spectrale : les sourdines des cuivres

On peut produire d'importants changements spectraux en introduisant dans le pavillon d'un instrument à embouchure différents objets appelés « sourdines ». Ce sont d'efficaces transformateurs de spectre qui, en modifiant certains modes propres, filtrent des zones du spectre et produisent des zones formantiques. Le Son **7.18** permet de comparer le son normal de la trompette à l'effet de trois sourdines : bol, sèche et wah-wah.

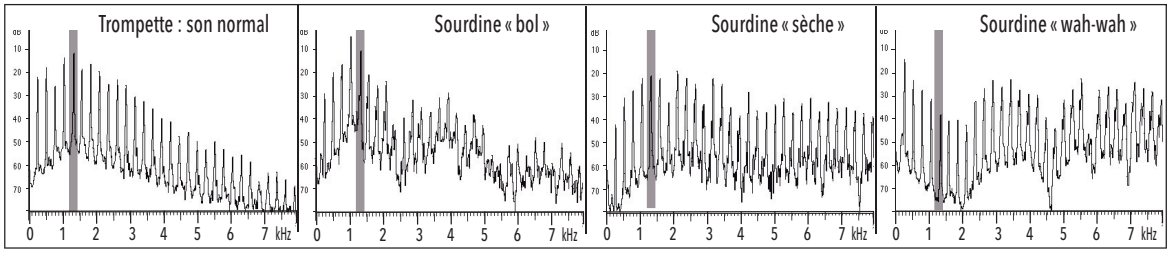


Figure 7.21 Analyse spectrale de l'effet de sourdines de trompette. Spectre LTAS de la note *do*$_3$ produite de façon ordinaire, puis avec trois sourdines différentes. (Base RWC ; analyses Wavesurfer ; échelle verticale de 80 dB, échelle horizontale de 0 à 8 kHz).

L'analyse spectrale moyennée de la note *do*$_3$ est donnée figure 7.21. Le formant habituel de la trompette, situé vers 1 300 Hz, est accentué par la sourdine bol qui « creuse » le spectre vers 2 500 Hz et atténue les harmoniques élevés ; il est conservé par la sourdine sèche qui renforce les aigus. C'est la sourdine « wah-wah » qui produit la transformation la plus radicale en inversant le spectre dans la zone des 1500 Hz et en donnant à l'enveloppe des harmoniques aigus un profil à plusieurs formants. Il n'est pas étonnant que les transformations opérées par les sourdines aient passionné le compositeur Gérard Grisey, pionnier de la musique spectrale, pour qui nous avons effectué une série d'analyses des différents modèles de sourdines existant pour la trompette, le cor et le trombone.

Son 7.18 (11″)

Un son de trompette et trois sourdines (note do$_3$*)*

Quelques caractéristiques remarquables du contenu spectral des sons de hauteur définie

1. L'enveloppe spectrale globale du spectre harmonique et la bande passante réduite (flûtes) ou large (anches). Prendre en compte les bruits, souvent délaissés comme le souffle, qui se situent souvent dans la partie haute du spectre.
2. Le rapport d'énergie entre le fondamental, ou harmonique 1, et l'ensemble du spectre.
3. La balance entre harmoniques pairs et harmoniques impairs.
4. Le caractère particulier de l'harmonique 5.
5. Les formants spectraux, qui sont fixes dans les instruments (voir le hautbois, le basson, les sourdines des « cuivres »). Leur interprétation en termes de voyelles doit être faite avec prudence.

Ces critères isolés sont rarement déterminants, c'est-à-dire qu'un seul d'entre eux n'est pas décisif. Ils prennent sens dans la comparaison, lorsqu'il faut décider de l'appartenance à une catégorie comme dans l'expérience de la « chimère » (voir §2.3.6.).

2.3.6. La clé d'écoute de la chimère

Nous avons maintenant tous les éléments pour comprendre l'identification du son présenté dans l'expérience du § 2.3.2, page 309 (Son **7.13**).

Le schéma de la figure 7.22 propose une interprétation temporelle (de haut en bas) de la séquence de décisions binaires conduisant à l'identification de l'instrument, en mettant en parallèle les événements sonores à l'origine du son et la navigation dans la hiérarchie des classes d'instruments.

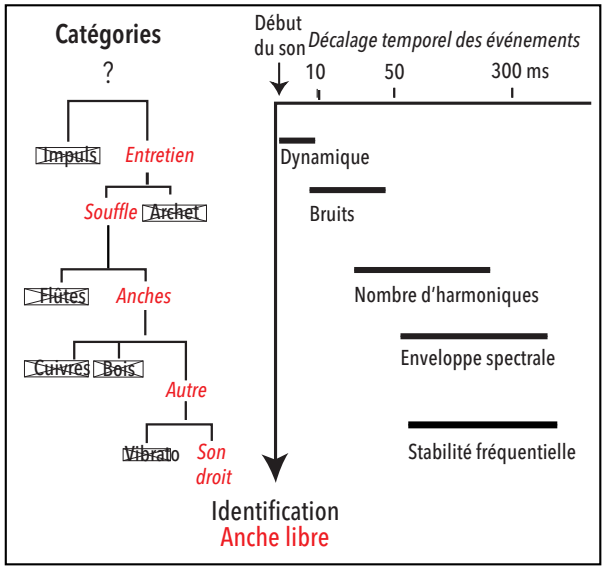


Figure 7.22 Exemple d'interprétation perceptive de la succession des choix de critères conduisant à la reconnaissance de l'instrument qui a vraisemblablement joué le son de l'exemple 7.13. Le temps se déroule du haut vers le bas de la figure. La partie gauche figure les bifurcations aux différents niveaux de catégorisation.

Le premier indice temporel, de nature dynamique, permet de décider du mode d'excitation : impulsion ou entretien. Dans notre exemple, le mode impulsion est rejeté. Nous avons ensuite deux possibilités d'entretien : l'archet et le souffle. La microstructure du bruit présent au début du son et le décalage temporel entre ce bruit et les premières composantes orientent sur la classe des instruments à vent. Si la durée relativement longue du transitoire d'attaque tendrait à orienter sur la flûte, celle-ci est éliminée en raison de la richesse en harmoniques et de la faiblesse du fondamental. Le choix se porte alors sur les instruments à anche. Le son étant dépourvu d'instabilité fréquentielle à l'attaque, les cuivres sont rarement mentionnés. Enfin, l'absence de formants et la présence d'harmoniques de tous rangs écartent les instruments de la famille du hautbois et de la clarinette.

Au bout du compte, il s'agit d'un son d'instrument à vent, riche en harmoniques, à transitoire d'attaque « mou » et qui, de surcroît, est dépourvu de vibrato. Ces caractéristiques correspondent aux sons de la classe des anches libres (accordéon, harmonica, harmonium), noms les plus fréquemment cités par les auditeurs. La structure réelle du Son **7.13** est donnée dans l'encadré.

Les premières millisecondes renseignent sur le geste effectué pour faire vibrer l'instrument (souffle, frottement) ; les caractéristiques spectrales qui suivent sont interprétées en cohérence avec les événements initiaux. La réunion des deux, qui constitue le transitoire proprement dit, est une entité, une forme sonore globale. À plus long terme, de l'ordre de la seconde, la reconnaissance s'affine progressivement sur la base de particularités relevant de spécificités instrumentales (stabilité fréquentielle) et stylistiques (vibrato).

Le son 7.13 : une chimère vraisemblable

Figure 7.23 Une chimère mythologique.
simlik/Shutterstock

À l'instar des chimères, gargouilles et autres monstres, le son **7.13** est un assemblage acoustique produit en laboratoire. Le plus étonnant dans cet exemple est sans doute le fait qu'il soit « vraisemblable » et donne lieu à des réponses logiques aboutissant à la désignation d'un instrument à anche alors qu'il est constitué de deux « portions » de son empruntées, d'une part, à un son de flûte (souffle + 90 ms du début du spectre) et, d'autre part, à un son de violon pour la partie stable du son (1,28 s).

Cette expérience montre bien que ni le transitoire d'attaque ni le contenu spectral du son pris isolément ne sont des indices déterminants. L'identité causale d'un son instrumental se construit sur la base d'une forme acoustique au sens « gestaltiste » du terme, c'est-à-dire une entité qui est autre chose que la somme des parties qui la constituent. On peut écouter dans l'exemple **7.19** le son de violon original ayant servi au montage.

Son 7.19 (3")

Son original

Il est important de signaler ici que les caractéristiques spectrales du son entretenu (richesse harmonique) ainsi que le caractère stable de la fréquence prennent sens au sein de la catégorie des instruments à vent, bien que la « matière acoustique » de la partie entretenue du son soit celle d'une corde frottée.

Cette expérience montre aussi que, du point de vue de la reconnaissance des sources, il est pertinent de distinguer les flûtes des anches, alors que ces deux groupes sont souvent rassemblés sous la dénomination globale de « bois ». Pour des raisons acoustiques, les sons des instruments à embouchure de flûte concentrent une grande partie de l'énergie dans la fréquence du premier harmonique (le fondamental) alors que ceux des instruments à anche offrent une répartition de l'énergie sur un grand nombre d'harmoniques. Il s'agit d'une tendance générale que l'on retrouve dans le vocabulaire : un son « flûté » s'oppose à un son « anché », du moins en Occident. Qui d'entre nous reconnaît une flûte dans l'exemple musical **7.20** ? La richesse harmonique et la faiblesse du fondamental incitent plutôt à penser qu'il s'agirait d'un hautbois. Or, c'est une flûte traversière dont le contenu spectral est transformé par la vibration d'une mince membrane qui fait office de mirliton, placée sur un orifice près de l'embouchure. Comme nous l'avons remarqué à maintes reprises, la catégorisation perceptive est une construction individuelle, valable pour une culture donnée. Celle dont nous traitons ici s'inscrit dans le cadre de la musique occidentale.

Son 7.20 (21″)

Voir page 387

2.4. Du son isolé à l'instrument : la variabilité de la forme identitaire avec la tessiture

Les caractéristiques identitaires que nous venons de décrire sont limitées aux deux notes considérées. Or, la plupart des instruments fonctionnent sur une étendue de deux à trois octaves. Le changement de fréquence implique des modifications de l'oscillateur et, dans certains cas, des structures couplées. La question se pose de la permanence ou des changements que subissent les formes spectrales que nous avons observées pour deux notes de l'octave 3. Prenons comme exemple initial l'orgue à tuyaux.

2.4.1. Deux cas particuliers : l'orgue et la voix humaine

Orgue

Parmi les premiers instruments de synthèse censés reproduire les sons des instruments naturels, certains fonctionnaient selon un principe élémentaire. Partant d'un spectre harmonique type, défini dans le medium, les notes supérieures et inférieures étaient obtenues par transposition de la fréquence avec la même enveloppe spectrale. C'est en quelque sorte le principe de l'orgue classique. Voici un exemple qui pose un problème perceptif intéressant.

Le Son **7.21** est l'enregistrement d'une gamme chromatique jouée sur un jeu de « tierce », en commençant par l'aigu du clavier. Curieusement, au fur et à mesure que la note descend, la perception se transforme.

Son 7.21 (62″)

Gamme chromatique descendante jouée sur le jeu de tierce d'un orgue (six rangs)

Nous découvrons avec surprise que la gamme chromatique commencée sur une note se termine sur un accord parfait majeur! Une deuxième écoute attentive nous fait prendre conscience de l'apparition progressive, au cours de la descente, de la quinte, puis de la tierce, qui s'affirment chacune comme note isolée dès que leur fréquence passe en dessous de 1000 Hz (contre-*ut*), donc pratiquement lorsque ces tuyaux entrent dans la tessiture de la clé de *sol*.

La composition spectrale du jeu reste inchangée, comme on peut le voir figure 7.25, et les intervalles entre harmoniques sont immuables. Il faut donc attribuer ces changements à la perception et en particulier à l'ambivalence hauteur/timbre qui varie selon la fréquence fondamentale. Plusieurs remarques vont dans ce sens. En premier lieu, les musiciens s'accordent pour dire que les timbres instrumentaux sont beaucoup moins différenciés dans l'aigu que dans le grave. En deuxième lieu, Helmholtz remarque à diverses reprises que les harmoniques sont difficiles à entendre isolément, mais qu'il est plus facile d'isoler auditivement les harmoniques impairs dont la note est différente de la fondamentale.

Le « jeu de tierce » de l'orgue classique

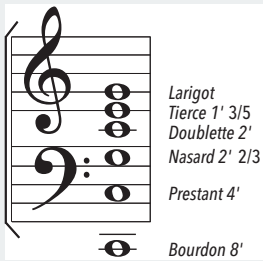


Figure 7.24 Jeux composant un jeu de tierce.

L'orgue est le premier et le plus important synthétiseur sonore dans lequel le timbre est, depuis longtemps, l'objet de la plus grande attention des facteurs. La caractérisation timbrale des jeux de flûte et de leurs combinaisons y est poussée à son plus haut niveau, avec un grand raffinement. Dans cet instrument, chaque tuyau ne produit qu'un son. Ceci permet d'ajuster les caractéristiques sonores de chaque note, indépendamment des voisines, et d'équilibrer en intensité et en sonorité la série des tuyaux d'un jeu donné sur plus de quatre octaves. Le terme « jeu » doit être pris ici dans le sens d'instrument. La position d'un jeu dans la tessiture est désignée par la longueur, en pieds, de son tuyau le plus grave. Les tuyaux d'un jeu de 8 pieds (8') sont à l'unisson des notes d'un piano. Ceux des jeux de 4'et 2', sont à intervalles d'octave. Quelques jeux particuliers de flûtes sont construits pour sonner à l'unisson de l'harmonique 3 (nasard 2'2/3)) ou de l'harmonique 5 (tierce 1'3/5) du tuyau de 8'. De grosse taille[1], ils n'émettent pratiquement que le son fondamental. En combinant plusieurs rangs de tuyaux on crée des synthèses harmoniques originales. Le jeu de tierce (figure ci-contre) associe cinq rangs de tuyaux[2] qui sont le bourdon, le prestant, le nasard, la doublette, la tierce (et, dans notre exemple, le larigot à l'octave du nasard). Cette combinaison est stable sur quatre octaves (voir le sonagramme de la figure 7.25), mais, pour l'auditeur, le caractère sonore du jeu se transforme de l'aigu au grave comme on peut le constater à l'écoute de l'exemple **7.21**.

Le *do* du milieu du clavier (environ 250 Hz) représente une sorte de frontière perceptive entre deux zones. La tessiture des deux octaves aiguës est comparable à celle du « cornet », autre jeu de l'orgue constitué de cinq rangées de tuyaux non séparables et s'arrêtant au milieu du clavier. Dans la partie supérieure du clavier, les deux jeux, cornet et jeu de tierce, peuvent dialoguer de concert. C'est la tessiture des deux octaves graves qui est caractéristique du jeu de tierce proprement dit. Du point de vue perceptif, l'aigu et le grave du jeu de tierce correspondent de fait à des timbres différents.

1. Sur la « taille » d'un tuyau, voir chapitre 2, § 4.6.
2. Notre exemple (orgue de la cathédrale de Poitiers) comprend six rangs. Noter le diapason très bas, et l'absence du premier *do*# grave.

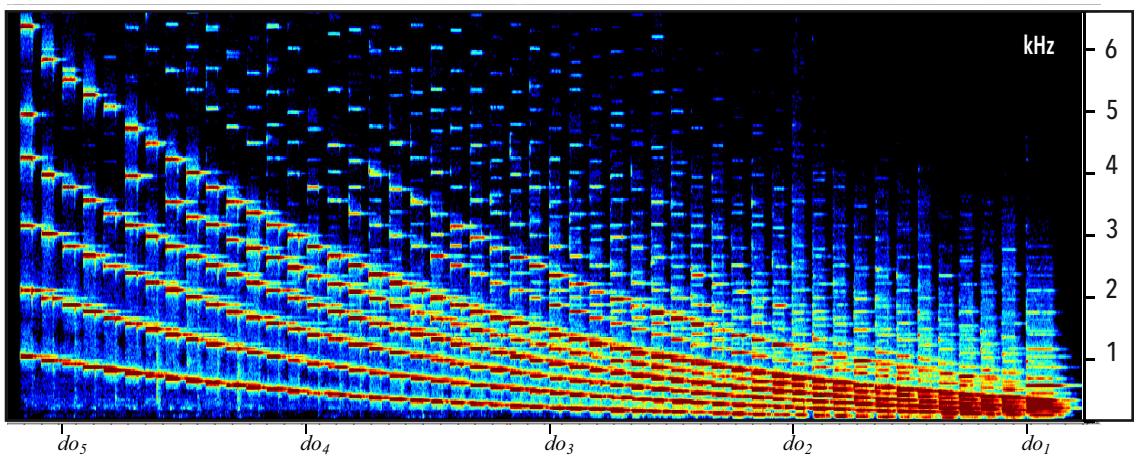


Figure 7.25 Analyse spectrographique du jeu de tierce (6 rangs) joué en gamme chromatique descendante. Clavier du Grand-Orgue de la cathédrale de Poitiers, avant restauration.

Son 7.21 (62")

Gamme chromatique descendante jouée sur le jeu de tierce d'un orgue (six rangs)

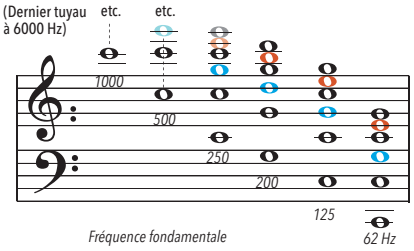


Figure 7.26 Interprétation des transformations de l'écoute à l'audition du Son 7.21.

Tentons une explication de notre expérience (figure 7.26). Dans l'aigu et le suraigu, de 2000 à 4000 Hz, le fondamental est dominant. Dans la zone médium aigu, celle du deuxième formant vocalique (1000-2500 Hz), les harmoniques 2 et 3 contribuent à renforcer la fréquence fondamentale qui reste dominante. Quand le fondamental passe de 500 à 250 Hz, l'harmonique 2 (octave) passe de 1000 à 500 Hz : il est entendu mais pas véritablement individualisé puisque c'est l'octave. Continuons à descendre : l'harmonique 3 (quinte, en bleu-vert) puis l'harmonique 5 (tierce, en rouge) entrent successivement dans la zone des fondamentaux vocaux et deviennent repérables. Pouvoir isoler l'une ou l'autre de ces fréquences dépend de l'attention des auditeurs et de leur capacité à se focaliser sur un harmonique. Pour beaucoup d'entre eux, l'harmonique 3 « sort » lorsque la fondamentale atteint le sol_3 ; l'harmonique 5 « sort » vers le la_2 dans cet exemple. La suite de la gamme chromatique descendante semble jouée en accords.

Voix humaine

À l'opposé d'un spectre de jeu d'orgue se translatant avec la fréquence se trouve la voix humaine pour laquelle la fréquence fondamentale varie indépendamment des fréquences des formants vocaliques, du moins dans l'ambitus de notre voix parlée ordinaire (voir chapitre 6, § 2.7). La translation en fréquence des formants de la voix humaine produit de curieux effets connus sous le nom d'effet « hélium »[39]. On peut le réaliser avec un logiciel qui permet de transposer l'enveloppe spectrale indépendamment de la fréquence fondamentale de la voix. Le Son **7.22** est un exemple réel issu du site de Joe Wolfe où l'on trouvera une analyse détaillée. Ce qui surprend dans cet exemple est moins l'anamorphose en fréquence des formants

Son 7.22 (10")

Effet « hélium » sur une voix d'homme *(voir page 388)*

39. Pour une explication claire et documentée voir le site de Joe Wolfe : http://phys.unsw.edu.au/phys_about/PHYSICS!/SPEECH_HELIUM/speech.html. Voir aussi le site de Kawahara dont le logiciel « STRAIGHT » permet de modifier aisément la fréquence fondamentale ou les formants.

(environ 1/3), qui varie dans la réalité selon l'âge et le sexe des personnes, que l'étrange coexistence de formants élevés avec une fréquence fondamentale grave.

La plupart des auditeurs notent que le « timbre change », ce qui montre bien que ce mot recouvre à la fois la perception de hauteur, la justesse des voyelles et même l'identité de la personne qui parle et son genre. Il est difficile de croire, à l'écoute de l'exemple, que la fréquence fondamentale de la voix est la même dans les deux exemples, ce qui est pourtant le cas.

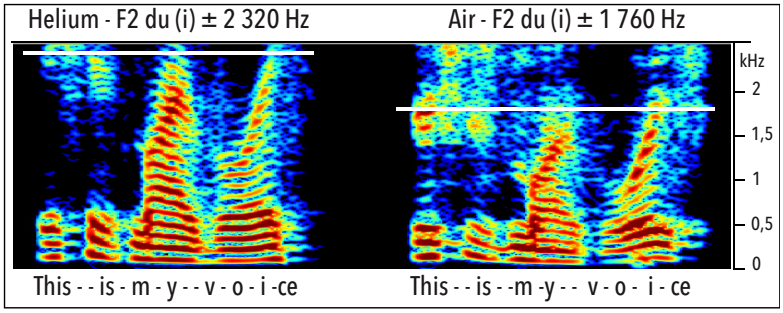


Figure 7.27 Sur l'analyse spectrographique, le trait rouge indique la position du deuxième formant du « i » qui revient à plusieurs reprises dans la phrase. L'effet de l'hélium (gauche) réalise une anamorphose fréquentielle des formants sans changer la fréquence fondamentale contrôlée par les cordes vocales.

Orgue et voix humaine sont des cas extrêmes. Que se passe-t-il dans une source instrumentale lorsque la fréquence fondamentale change ?

2.4.2. Les instruments de l'orchestre

Les modifications du contenu spectral avec la fréquence d'émission sont dues à la fois aux transformations mécaniques nécessitées par le changement de fréquence et aux modifications du rayonnement des sources : elles diffèrent donc selon les instruments. Dans les instruments où chaque note est produite par un oscillateur plus ou moins indépendant (cordes libres, orgue à tuyau, accordéon), le facteur a accès à l'harmonisation de chacune séparément, en équilibrant l'homogénéité selon son goût. Dans la plupart des instruments à vent, il n'y a qu'un oscillateur pour toute l'étendue. Les cordes à manche et les cordes frottées en possèdent plusieurs : 4 le plus souvent, jusqu'à 7 ou 8.

Cordes

Le changement de fréquence d'un instrument à cordes s'obtient par raccourcissement de longueur, ce qui, pour une position donnée de l'excitateur, modifie peu l'enveloppe globale du son, dans la mesure où les résonances de la caisse sont réparties de façon assez homogène. En revanche, le rayonnement, extrêmement complexe, se modifie en permanence. Ces instruments n'ont pas de registre à proprement parler. Le spectre vit, se transforme d'une note à l'autre sans subir de changement majeur entre le grave et l'aigu, excepté pour les sons voisins de la résonance d'Helmholtz, généralement située dans le grave de la tessiture (aux alentours de 280 Hz pour le violon, soit $ré_3$).

Vents

Dans la classe des instruments à vent, il faut distinguer deux groupes : les instruments à embouchure et les instruments à trous latéraux.

Instruments à embouchure

Dans les instruments à embouchure de longueur fixe (clairon, cor des Alpes, trompe de chasse), le changement de hauteur fondamentale provient des lèvres qui choisissent différents partiels du tube. De tels instruments ne peuvent jouer qu'une suite discontinue de sons. Les instruments d'orchestre possèdent des rallonges qui permettent de transposer la fréquence fondamentale. L'enveloppe spectrale des instruments à embouchure reste quasi constante sur toute l'étendue, chaque famille d'instruments possédant un formant spécifique (Luce & Clarke, 1967). Les « cuivres » rayonnent le son par une ouverture unique ce qui, dans une certaine mesure, contribue à leur reconnaissance en audition directe, en particulier à l'orchestre. Les harmoniques aigus, prépondérants dans l'axe du pavillon, sont très présents dans la trompette et le trombone qui sont dirigés vers les auditeurs, mais plus atténués dans le cor dont le pavillon est tourné vers l'arrière. Quand au tuba, le pavillon est habituellement dirigé vers le plafond.

Instruments à trous latéraux

Le changement de fréquence des instruments à trous latéraux s'obtient en modifiant les vibrations internes de l'air enclos dans le tube, par l'ouverture ou la fermeture de trous. Le nombre des doigts étant nécessairement limité, ces instruments ont recours aussi aux partiels des modes supérieurs. Certains octavient, d'autres quintoient. En conséquence, les sons produits sur la base du partiel 1 n'ont pas le même contenu spectral que ceux qui dépassent l'octave (flûte, hautbois, saxophone) ou la douzième (clarinette).

La plupart des instruments à trous latéraux possèdent des « registres » ou zones de la tessiture de timbre spectral similaires, qui sont indiqués dans les traités d'orchestration. L'instrumentiste a peu de possibilités pour les modifier : ce sont des traits identitaires de l'instrument. Dans l'aigu, il dispose souvent de plusieurs doigtés différents pour produire une note donnée, ce qui lui procure une certaine marge de variation.

La variabilité spectrale principale provient des changements incessants du nombre et de la position des trous ouverts, ce qui conditionne le spectre des notes jouées et a des incidences complexes sur le rayonnement. Ceci explique en partie la difficulté d'identification de ces instruments, exceptée la clarinette dont les premiers sons graves (les 16 sons du registre de « chalumeau ») sont très caractéristiques. A contrario, ce sont justement les inhomogénéités spectrales dues aux sons de fourche qui contribuent à la reconnaissance des instruments baroques (voir Castellengo, Drouin, Séchet, 1978 ; Castellengo & Forest, 1999).

2.4.3. Un survol spectral des principales catégories de sons entretenus : les gammes chromatiques en jeu lié et détaché

Jeu *legato*

Les figures 7.28 et 7.29 permettent de comparer les analyses de gammes chromatiques jouées sur les sept sortes d'instruments mélodiques que nous avons examinés plus particulièrement et auxquels nous avons ajouté le basson. Jouer une gamme chromatique rapidement, sans erreur, avec justesse et égalité de son est un exercice difficile, quasi automatique, que les musiciens professionnels pratiquent quotidiennement pour entretenir la maîtrise coordonnée des doigtés et de l'émission sonore. L'expressivité en est quasiment absente. Certains musiciens y sont d'ailleurs tellement rompus que la reproductibilité spectrale de leur prestation est étonnante.

Les huit analyses que nous présentons figures 7.28 et 7.29 correspondent à des enregistrements effectués au LAM, de 1965 à 2014. Ces enregistrements portent la trace des modèles d'instruments utilisés ainsi que celle de l'idéal sonore des musiciens qui ont accepté de jouer pour nous. Il ne faut donc pas chercher à en extraire des indications précises sur une note particulière d'autant qu'on sait que le spectre capté dépend aussi de la position du microphone et de l'acoustique du local d'enregistrement. Mais, dans l'objectif que nous nous sommes donné, qui est de rendre compte de la caractérisation sonore des principales familles instrumentales, nous pouvons considérer qu'un enregistrement de gamme chromatique donne un bon aperçu d'ensemble de la répartition spectrale des sons d'un instrument, et en particulier des changements avec la tessiture.

À première vue, chaque instrument présente une allure caractéristique montrant une certaine homogénéité et des traits spécifiques qui lui sont propres.

Bande passante. L'examen du spectre harmonique sur l'ensemble de la tessiture permet de compléter la caractérisation des sources esquissée au § 2.3.5 (caractérisation acoustique de la note). Deux instruments : le violon et l'accordéon couvrent la totalité du spectre analysé pour tous les sons émis (et au-delà : jusqu'à 14 kHz). Le hautbois et le basson couvrent une zone spectrale constante limitée à 7 kHz pour le premier et à 3 kHz pour le second. Pour les autres instruments : clarinette, saxophone, flûte traversière et trompette, la bande passante croît avec la tessiture.

Registres. Nous avons indiqué sur les deux figures les zones correspondant aux registres des instruments à trous latéraux. La zone noire correspond à celle du premier registre : zone des fondamentaux du tube ou « régime 1 ». La zone rouge correspond aux sons du deuxième registre : zone du partiel 2 ou « régime 2 », octaviant pour la flûte, le hautbois, le saxophone, le basson, et quintoyant pour la clarinette.

Voici quelques observations spécifiques. Remarquez que les notes écrites correspondent aux sons réels et que la note la_3 est indiquée en rouge sur toutes les analyses.

Clarinette : notez la faiblesse des harmoniques pairs dans le premier registre et leur renforcement à partir du sol_3.

Basson (modèle français) : on note une large zone formantique (300-600 Hz), une plus étroite (1400 Hz) et deux zones blanches ou anti-formants vers 900 et 2050 Hz.

Hautbois : le formant signalé au § 2.3.5 est bien visible aussi.

Flûte traversière : les notes $ré_4$ et $ré_5$ délimitent trois zones spectrales (voir les traces chuchotées entre les harmoniques). Le fondamental acoustique du tube est $ré_3$: les trous de la patte (rallonge do_3-$do\#_3$) restent ouverts.

Saxophone alto : en régime 1, on note la prédominance de l'harmonique 2 et souvent l'octaviation de l'attaque ; dès le fa_3, le fondamental et l'harmonique 3 se renforcent (visible aussi en jeu détaché, voir figure 7.30).

Trompette : elle ne présente pas de registre à proprement parler. La zone formantique s'étendant de 1 à 3 kHz, repérée sur les spectres de la figure 7.18, est homogène sur toute la tessiture. Remarquez les notes graves, très riches en harmoniques, dites « cuivrées ».

Son 7.23a (4,8″)

Hautbois (Debray)

Son 7.23b (7″)

Clarinette (Ettlinger)

Son 7.23c (3,7″)

Saxophone (Delangle)

Son 7.23d (4,2″)

Basson, modèle Buffet (Allard)

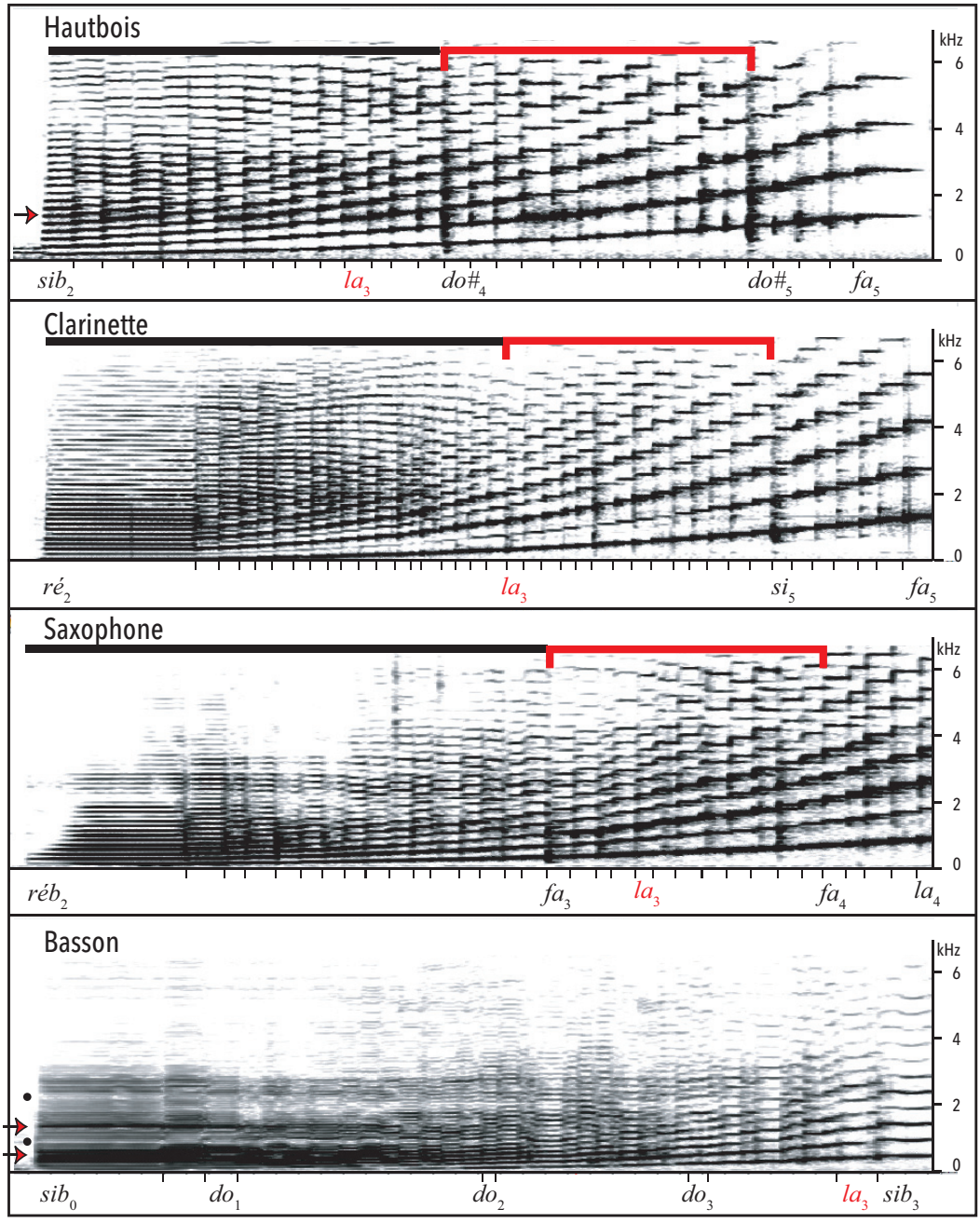
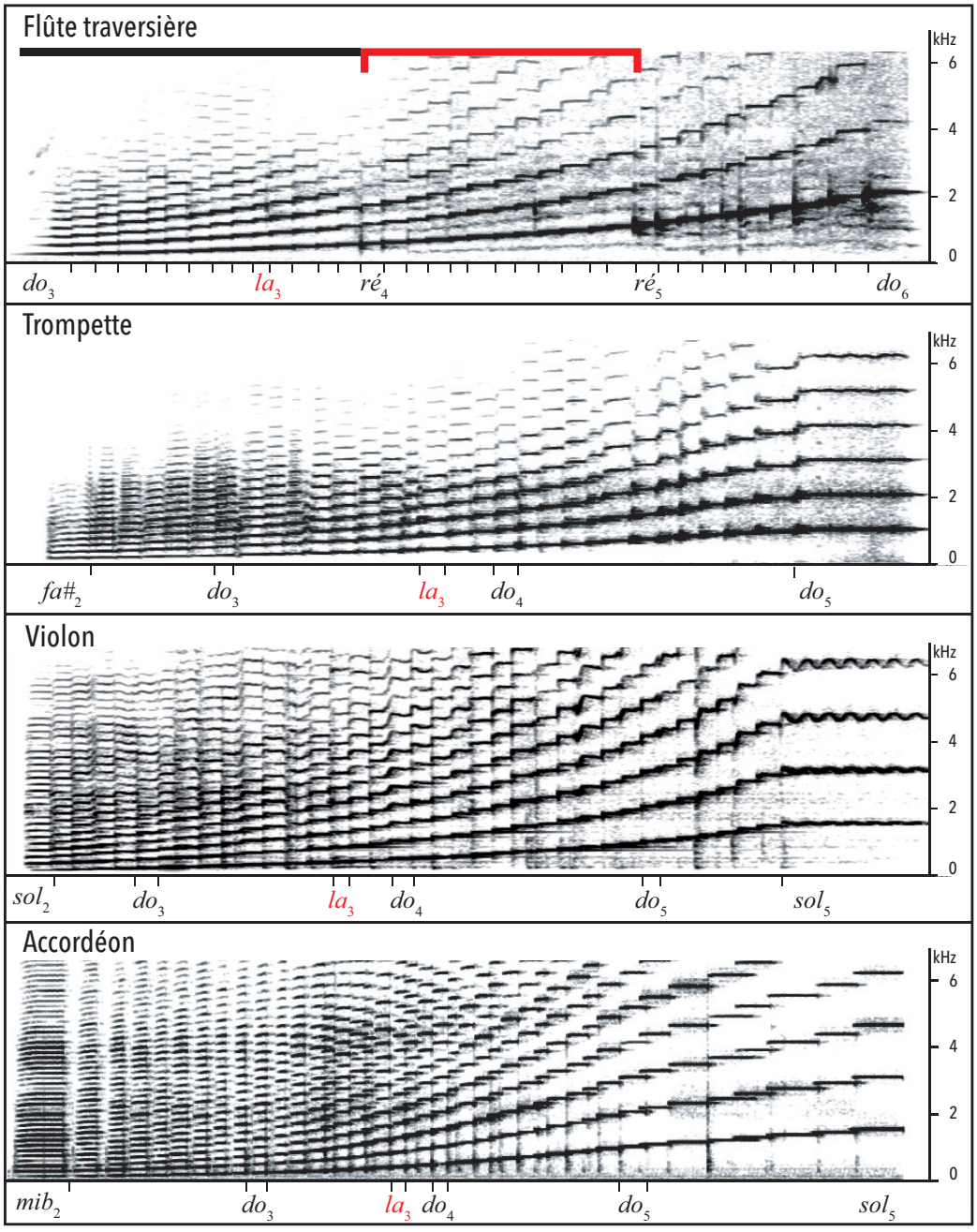


Figure 7.28 Analyses spectrographiques de gammes chromatiques jouées *legato* ; nuance *mf*. Les noms des notes sont en hauteur réelle. La durée indiquée est celle de l'extrait analysé. Les traits de couleur noire et rouge figurent l'étendue du 1er et du 2e registre. Les flèches rouges indiquent les zones formantiques et les points les zones de réjection ou antiformants. Analyses : FFT 2 048 pts ; échelle de fréquence linéaire, maximum 7 kHz.



Son 7.23e (5,5″)

Flûte traversière (Fabre)

Son 7.23f (5,5″)

Trompette (Roquin)

Son 7.23g (7″)

Violon (Gill)

Son 7.23h (8″)

Accordéon (Monichon)

Figure 7.29 Analyses spectrographiques de gammes chromatiques jouées *legato* ; nuance *mf*. Les noms des notes sont en hauteur réelle. La durée indiquée est celle de l'extrait analysé. Les traits de couleur noire et rouge figurent l'étendue du 1er et du 2e registre. Analyses : FFT 2 048 pts ; échelle de fréquence linéaire, maximum 7 kHz.

Jeu détaché

Voici maintenant l'analyse des gammes chromatiques jouées en jeu détaché (figure 7.30). Ce type d'articulation rend mieux compte du jeu ordinaire de la trompette et accentue les phénomènes se produisant à l'attaque (voir notamment les bruits du coup de langue à la flûte et ceux de l'archet de violon). Noter en particulier le renforcement des partiels des régimes inférieurs à l'attaque et pendant le son, sous forme de bruits colorés : flûte traversière à partir du *ré*$_4$, saxophone à partir du *fa*$_3$. Pour ce dernier on peut également noter un autre changement spectral : l'harmonique 3 est plus marqué dans le deuxième registre.

Gammes chromatiques ascendantes en jeu détaché

Son 7.24a (10")

Flûte (Artaud)

Son 7.24b (19")

Saxophone, gamme ascendante descendante (Delangle)

Son 7.24c (8")

Trompette (Roquin)

Son 7.24d (10")

Violon (Gill)

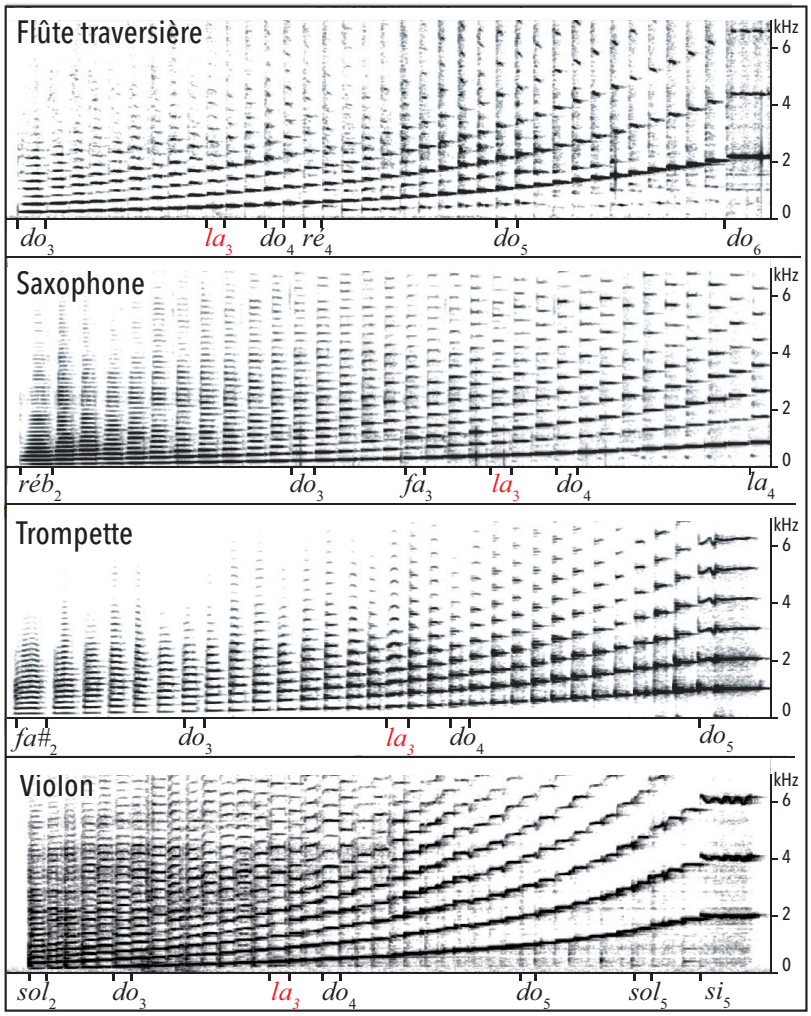


Figure 7.30 Analyses de gammes chromatiques en jeu détaché : flûte traversière, saxophone, trompette, violon. Notes réelles. FFT : 2 048 pts. Échelle de fréquence 7 kHz ; échelle de temps variable.

Les traits particuliers que nous signalons ici ne dépendent que de la perce et du système de cléterie de l'instrument. Ils sont indépendants du musicien pour peu que celui-ci utilise les doigtés de base courants, ce qui est généralement le cas lors du jeu d'une gamme chromatique rapide.

2.4.4. La variabilité des transitoires avec la tessiture

Dans la plupart des techniques de synthèse, le son est découpé en trois parties : le début ou transitoire d'attaque, la partie tenue et la fin ou transitoire d'extinction. En acoustique, le repérage de la partie transitoire est souvent défini par la pente de l'enveloppe d'amplitude. Or, celle-ci dépend, entre autres, de la tessiture de jeu. Un corps plus gros et plus lourd met plus de temps à se mettre en vibration qu'un autre plus petit et plus léger, si bien que, d'une façon générale, la durée minimale des transitoires d'attaque diminue du grave à l'aigu, qu'il s'agisse des variations au sein d'un instrument donné ou entre les instruments d'une même famille (flûte en *ut* et flûte basse, par exemple).

Certaines sources instrumentales comme les anches doubles (hautbois, basson) et les cordes frottées, présentent peu d'inertie : le son démarre nettement et rapidement ; d'autres (flûtes, anches simples et embouchures) sont fortement dépendantes des dimensions géométriques de l'instrument.

À titre d'exemple, voici, pour quatre des instruments déjà examinés, une analyse du son le plus grave et du son le plus aigu de chacun d'entre eux. Compte tenu de la variabilité du jeu, nous avons sélectionné dans les bases de données les sons dont l'émission était la plus brève et la plus nette. Sur la figure 7.31, on voit successivement, de gauche à droite et de haut en bas (en notes réelles) : trompette ($fa\#_2$-do_5), accordéon (fa_2-la_5), violon (sol_2-do_5), clarinette ($ré_2$-fa_5).

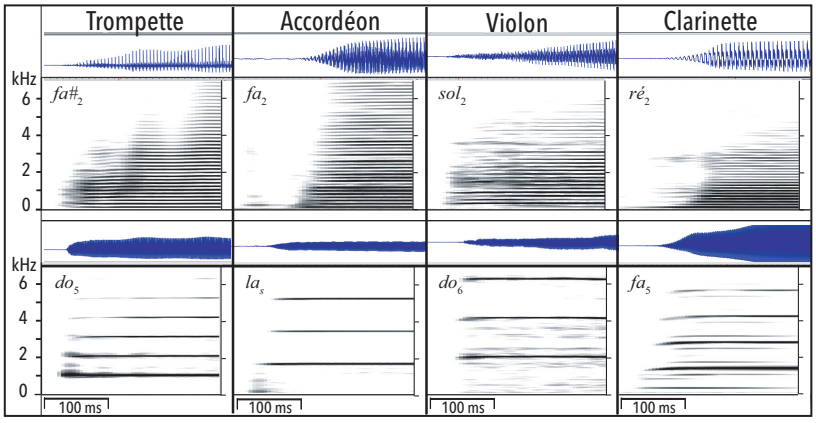


Figure 7.31 Variabilité du transitoire de quatre instruments en fonction de la fréquence d'émission (son le plus grave et son le plus aigu ; notes réelles). Échelle de temps 30 ms ; échelle de fréquence 7 kHz (2 048 pts). Références des sons : trompette et violon – sons de la base McGill ; accordéon et clarinette – sons de la base RWC (071ACR3F et 311CLNOM).

On considère habituellement que le transitoire d'attaque est la partie pendant laquelle se produisent les bruits de l'excitation (grincement, souffle, etc.) suivis par l'apparition des harmoniques en ordre variable. D'autres événements se produisent et notamment des fréquences étrangères à la série harmonique, des oscillations de la fréquence fondamentale : autant d'indices qui trahissent le mode de mise en route d'un système complexe constitué d'un excitateur et d'un résonateur possédant des modes propres sans rapport harmonique.

2.4.5. La variabilité sonore : un inventaire à faire

Il manque un inventaire exhaustif de la « flore et de la faune » sonores instrumentales. La première grande base de données des sons d'instruments de musique réalisée à l'Université de McGill (Opolko & Wapnick, 1987) offre plus de 6000 sons enregistrés avec les principaux instruments de l'orchestre, des instruments baroques et la plupart des instruments à percussion courants à l'époque. Elle a servi de fonds commun pour de nombreuses expériences de psychoacoustique et reste une référence pour la qualité musicale du jeu et des enregistrements. On ne dispose cependant que d'un son par instrument, avec une seule nuance de jeu.

Cette lacune est comblée par la base de l'université de l'Iowa (voir Bibliographie, rubrique « bases de données ») et par celle de RWC (Goto & coll. 2003) dans laquelle deux à trois musiciens accomplissent le programme d'enregistrement d'un instrument donné avec des modèles différents, ceci pour trois nuances de jeu. On y trouve aussi les instruments spécifiques de la musique japonaise. Pour les besoins de la recherche, l'Ircam a procédé à un inventaire des modes de jeu contemporains. Un problème demeure : ces entreprises considérables[40] sont destinées aux « sampleurs ». Elles offrent des sons isolés, dont nous avons fait largement usage, mais les transitions entre sons jouées en *legato* – sons conjoints ou surtout sons disjoints de registres différents –, tous les enchaînements qui font la richesse et la complexité d'un phrasé caractéristique du jeu instrumental en sont exclus. Ce travail fondamental reste à faire.

2.5. L'instrument de musique : permanence-variation

> *Tout dispositif qui permet d'obtenir une collection variée d'objets sonores – ou des objets sonores variés –, tout en maintenant présente à l'esprit la permanence d'une cause, est un instrument de musique, au sens traditionnel d'une expérience commune à toutes les civilisations.*
>
> Schaeffer, P., 1966, p. 51.

Expérience de P. Schaeffer (GRM)

Son 7.25a (35")

Sons modifiés

Son 7.25b (31")

Sons originaux

Permanence de la cause et variété des objets sonores, nous ne saurions mieux définir ce qu'est un instrument de musique. On sait les difficultés rencontrées par les musiciens qui créent les sons en studio. S'il est devenu assez aisé de produire des sons intéressants, il est souvent difficile de suggérer un semblant de causalité analogue à celui d'un instrument mécanique. Au cours de ses nombreuses explorations sonores, P. Schaeffer a réalisé une expérience préliminaire assez convaincante qui démontre l'incidence de l'enveloppe d'amplitude initiale sur la reconnaissance des sources.

Les huit sons que nous entendons au début de l'exemple **7.25a** semblent bien être produits par la même source acoustique alors que l'auteur a enregistré successivement les sons suivants : flûte, basson, flûte, clarinette, son de synthèse, hautbois, trompette, son de synthèse. Il a suffit d'affecter à chacun de ces sons la même enveloppe dynamique temporelle pour que nous acceptions d'en faire une collection homogène, une catégorie instrumentale.

L'enveloppe est constituée d'une croissance douce de l'amplitude pendant 1 s, suivie immédiatement d'une décroissance lente pendant 1,7 s. On peut entendre les sons originaux dans l'exemple **7.25b**.

40. Rien que pour le piano, la base Japonaise RWC comprend 3168 sons (3 instruments, 4 modes de jeu, 3 nuances d'intensité).

2.5.1. L'interprète et son instrument

Un instrumentiste « traditionnel » ne cesse de varier les sons qu'il produit, en particulier aux moments-clés de l'écoute que sont les transitions : début des sons et enchaînements des sons successifs. Dans le groupe des instruments mélodiques, il est sans conteste le maître de la qualité sonore[41]. Mais comme toutes les variations ne sont pas possibles sur tous les types d'instruments, certaines agissent comme critères déterminants de l'identité de la source.

Son 7.26a (5")

De quel instrument s'agit-il ?

Voici encore une petite expérience qui fait entendre une note isolée (Son **7.26a**). De quel instrument s'agit-il ? Une guitare ? un piano ? Oui et non. C'est peut-être un piano, car la percussion et la texture inharmonique du son avec ses battements sont caractéristiques de cet instrument. Non, car le son d'un piano ne peut être vibré. Un des critères-clé de la catégorisation des sons « piano » que chacun de nous élabore se fonde justement sur la stabilité fréquentielle. Le vibrato agit ici comme critère identitaire et suffit à exclure ce son de la catégorie « piano » alors qu'il s'agit d'un son réel de piano dont la fréquence fondamentale a été modulée.

Son 7.26b (5")

Son original

Vibrato

Le vibrato est possible aux instruments à cordes et aux instruments à vent (anche libre exceptée). Dans la mesure où sa réalisation varie selon les sources, il fait aussi partie des caractéristiques identitaires et complète celles que nous avons déjà définies. Sur certaines sources, le vibrato est principalement une variation de fréquence (cordes, voix), sur d'autres, une variation d'amplitude (saxophone jazz) ou une modulation combinant l'amplitude et la fréquence (flûte). Le vibrato est aussi un marqueur de style musical au point que la clarinette, qui a connu un grand engouement à une époque où le vibrato était peu répandu à l'orchestre, voire inexistant, est très reconnaissable dans le pupitre des « bois » par la stabilité de ses sons.

Nuances de jeu

Les nuances d'intensité déjà traitées au chapitre 5, sont également sous le contrôle de l'interprète. Nous avons vu qu'un crescendo est plus complexe qu'une seule augmentation de l'amplitude puisqu'il s'accompagne toujours d'un enrichissement en harmoniques, particulièrement dans la zone sensible de l'oreille, de modifications de l'attaque des sons et de changements dans le vibrato qui devient plus ample et un peu plus rapide. Avant même d'émettre le premier son, un instrumentiste entend déjà ce qu'il va produire : un son fort, faible, agressif ou doux, intentionnellement paré de qualités expressives. Penser le jeu d'un instrument est un geste global de l'ensemble du corps, du souffle ou de l'archet, qui contrôle la mise en vibration et la conduite du son, et c'est au travers de sa réalisation acoustique que les auditeurs perçoivent l'intentionnalité expressive de l'interprète (voir en particulier l'exemple du violon, chapitre 5, § 5.1).

2.5.2. Les transitoires d'attaque et le style de jeu de l'interprète : sons des bases de données

Les variabilités de l'attaque sont déjà présentes dans les bases de données qui offrent en principe des sons « neutres ». Les analyses de la figure 7.32 permettent de comparer deux instruments représentatifs de la variabilité due au style de jeu de l'interprète : le violon et la flûte traversière. Les sons de l'arpège do_3, mi_3, la_3, do_4, sont joués par deux musiciens différents, enregistrés dans des conditions similaires (en studio). Ces exemples correspondent à une « bonne » attaque, propre et précise. Toutes les analyses sont faites avec la même transformée de Fourier et sont présentées dans les mêmes échelles de fréquence (8 kHz) et de temps (300 ms) ; FFT : 1 024 pts.

41. À quelques exceptions près, en particulier lorsque le son est « fait » par le facteur : flûte à bec, accordéon et, bien sûr, l'orgue.

L'observation des sons disposés en colonnes confirme l'empreinte de l'interprète dans la production du son.

Le violoniste 1 a une attaque nette et précise alors que le violoniste 2 amène le son plus progressivement. Pour la flûte traversière, l'attaque du joueur 1 est constituée principalement de souffle diffus, alors que celle de 2 montre un coup de langue net qui s'accompagne de petits bruits graves et de sons de biseau ascendants. On notera aussi à l'écoute les différences dans le vibrato (non visibles sur l'analyse qui ne montre que le début du son). Flûte 1 : vibrato d'amplitude très marqué et très lent (4,5/s) ; flûte 2 : vibrato plus rapide (5,5 à 6/s) et moins ample.

Arpèges ascendants. Lire la figure correspondante de bas en haut

Son 7.27a (19")

Violon 1

Son 7.27b (14")

Violon 2

Son 7.27c (14")

Flûte traversière 1

Son 7.27d (19")

Flûte traversière 2

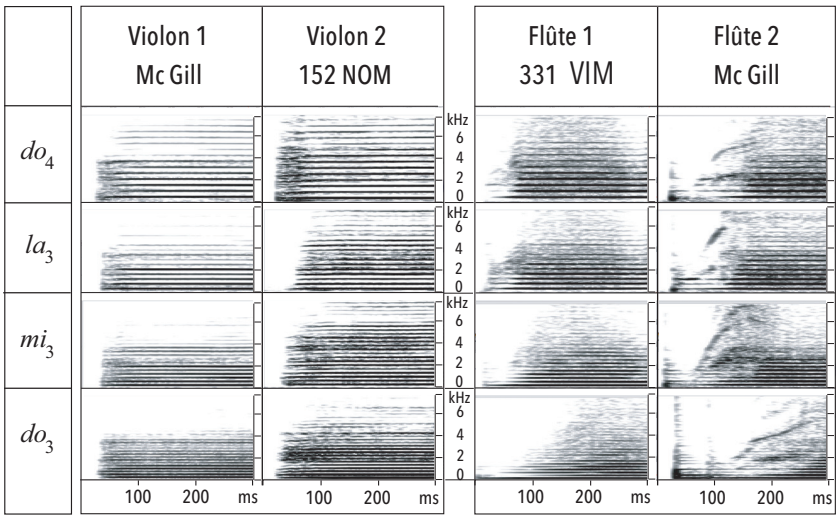
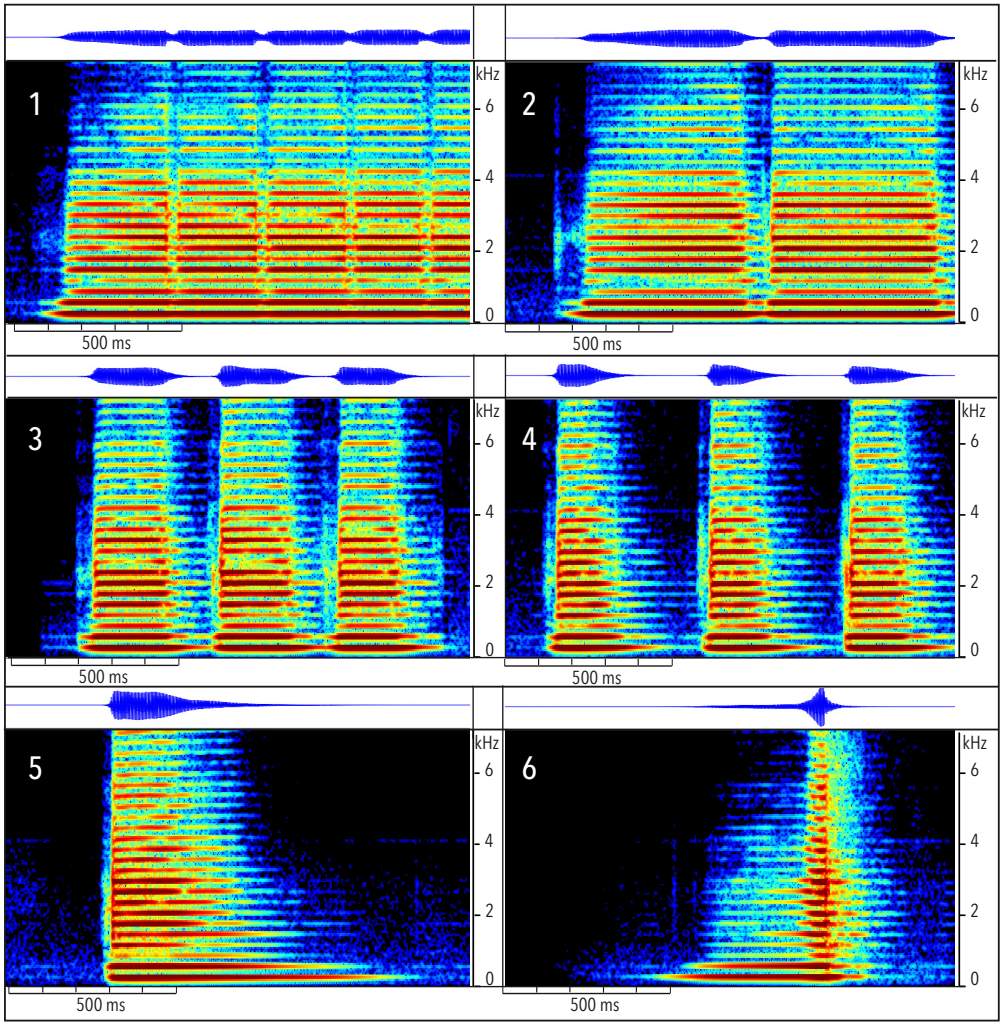


Figure 7.32 Analyses spectrographiques des parties transitoires de deux instruments à son entretenu : le violon et la flûte, joués par deux instrumentistes différents. Lire verticalement, de bas en haut, les quatre notes de l'arpège de *la* mineur (do_3 à do_4). (Bases de données RWC et McGill)

2.5.3. La conduite du son entretenu, de l'attaque à l'extinction : l'art du musicien (saxophone)

De tous temps, l'art de l'articulation des sons dans le jeu des instruments à vent emprunte aux mouvements de langue dans la prononciation des consonnes : T, D, K, L. Tout entre en ligne de compte pour la perception : la rapidité d'établissement du son, la netteté du petit bruit de souffle qui le précède mais aussi et surtout la façon dont le son est relâché.

La figure 7.33 présente quelques exemples joués au saxophone sur le $ré_3$ (Son **7.28**). Dans l'enchaînement de plusieurs notes, l'instrument peut être réattaqué sur le son (exemples 1, 2 et 3) ou après l'arrêt total (exemple 4). Dans ce cas, la perception n'en est que plus brusque, car la durée du silence d'interruption contribue aussi au caractère de l'attaque.



Son 7.28 (24")

Saxophone : attaques variées *(voir page 388)*

Figure 7.33 Exemples d'attaques variées sur une note de saxophone alto, *ré*$_3$, hauteur réelle. Articulation douce « de » sans interruption du son (1 et 2) : articulation précise « te » sans interruption (3), puis jeu détaché (4) ; attaque « cloche » (5) et attaque « cloche inversée » (6). Échelle de fréquence 7 kHz (1 024 pts) ; 1,4 s à l'écran. (C. Delangle).

Une attaque très brusque suivie d'un relâchement soutenu et décroissant simule la résonance d'une cloche (exemple 5) et le mouvement inverse (attaque très progressive se terminant par un arrêt soudain) évoque le renversement temporel du même son sur bande (exemple 6). Ces exemples montrent aussi les limites d'une étude fondée sur des sons isolés extraits de bases de données. Hors de tout contexte musical, les instrumentistes émettent des sons « neutres », sans qualité particulière. Or, l'expressivité musicale du jeu des instruments s'exerce principalement dans l'articulation des sons (successifs ou simultanés). Le mode d'attaque, les transitions liées ou détachées combinées au vibrato lorsqu'il existe et aux micronuances des transitions produisent des modifications du timbre instrumental perçu où se mélangent des indices spécifiques d'un instrument donné, auxquels sont sensibles les musiciens avertis, et les traits du style de jeu d'un interprète particulier. Caractériser un instrument nécessite d'explorer toutes les variations possibles de l'instrument, ses « champs de liberté » (Leipp, 1965, *Bulletin du GAM* n° 10).

2.5.4. Le son de flûte : entre son périodique et bruit de turbulence

De tous les instruments à vent, la flûte est sans doute celui qui offre la plus grande variété de modes d'attaque du son. Paradoxalement, le timbre flûte est une catégorie possédant à la fois une forte identité timbrale et une extrême diversité de timbres. À l'écoute d'une seule note, exercice d'écoute inframusical qui permet de s'abstraire des caractéristiques liées au style de musique, on remarque immédiatement les points communs aux sons de flûte (Son **7.29**) : la prédominance des premiers harmoniques et la présence permanente du souffle. Selon que ces flûtes servent des musiques de notes ou de timbre le dosage réciproque harmoniques/souffle varie de façon considérable comme on peut le voir sur les analyses de la figure 7.34 :

Son 7.29 (17")

Le même son (la_3) joué sur quatre instruments de la famille des flûtes. Lire la figure 7.34 par colonnes.

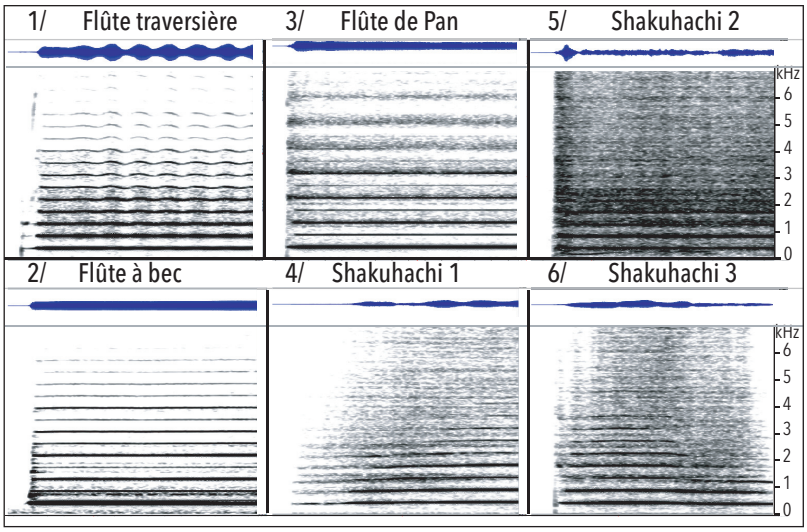


Figure 7.34 Six exemples extrêmement divers de timbre flûte, note la_3, durée 1,5 s. Analyses FFT de 2 048 pts ; 7 kHz ; 1,5 s. Sons extraits de bases de données 1 à 3, base McGill ; 4 à 6 base RWC.

1/ *Flûte traversière* : attaque nette (coup de langue avec petits bruits de bouche), nombreux harmoniques, souffle faible entre les harmoniques graves accentué par le vibrato. 2/ *Flûte à bec* : attaque très nette avec partiels inharmoniques, fondamental très intense, souffle peu audible, son très stable. 3/ *Flûte de Pan :* attaque de son et de souffle, ce dernier restant très présent tout au long du son dans la partie supérieure du spectre, au détriment des harmoniques ; forte prédominance des impairs (tuyau bouché). 4/, 5/, 6/ *Shakuhachi* : les trois exemples retenus, d'une très grande diversité dans l'utilisation du souffle, ne représentent qu'une petite partie des variations possibles de la production du son, tant à l'attaque qu'en phase finale (mélange son-souffle et attaques contrastées).

L'étude acoustique et cognitive du timbre de sons isolés est déjà très complexe. Il faudrait entreprendre celle du jeu des instruments dans différentes musiques, ce qui ne peut se faire qu'avec la participation des musiciens et des musicologues.

2.5.5. Les nouveaux modes de jeu en musique contemporaine

Tout objet mis en œuvre par un compositeur est un instrument de musique.

Berlioz, H., 1855, p. 2.

Cette surprenante définition de l'instrument de musique qu'Hector Berlioz place en tête de son ouvrage nous rappelle le lien étroit entre musique, instrument et composition. Berlioz publie le premier traité d'instrumentation au moment où la plupart des instruments subissent de profondes transformations afin de répondre au nouvel art sonore de l'orchestre qui exige justesse, puissance et homogénéité de sonorité d'un instrument, sur toute sa tessiture. Au milieu du XXe siècle, la musique se transforme. La quête de nouvelles sonorités conduit les compositeurs qui écrivent pour les instruments, à contrecarrer les codes habituels du « beau son ». L'accent est mis sur les bruits, sur l'inouï.

Son 7.30 (15")

Flûte traversière, les modes de jeu successifs analysés figure 7.35

À l'écoute du Son **7.30**, il est difficile de croire que tout ce que nous entendons « est » de la flûte traversière!

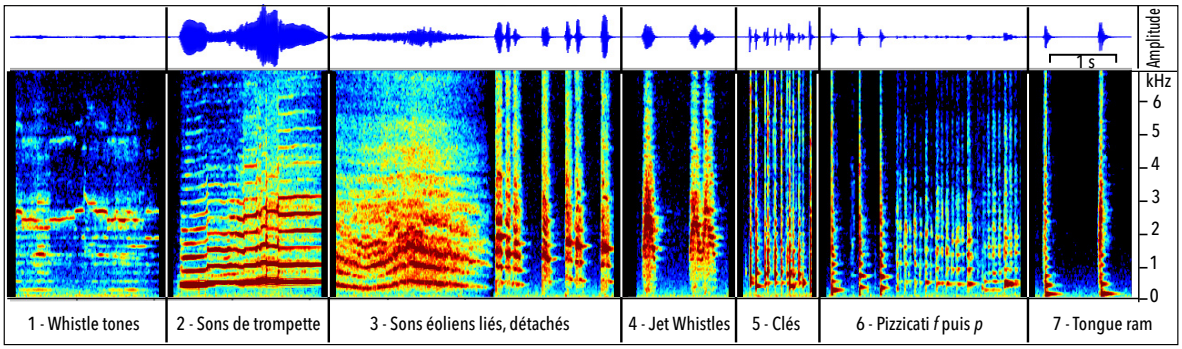


Figure 7.35 Analyse de modes de jeu de la flûte traversière contemporaine (P.-Y. Artaud).

Une telle diversité provoque l'éclatement perceptif des catégories traditionnelles : un son percuté ne « peut pas » être de la flûte. Sous l'impulsion de compositeurs comme Varèse, Jolivet, Mefano ou Levinas, la flûte traversière a été pionnière des nouveaux modes de jeu que de grands solistes (René Le Roy, P.-Y. Artaud) se sont appropriés (Artaud & Geay, 1980). En premier lieu, citons l'exploitation du souffle, élément qui fait partie intégrante du jeu de la flûte dans la plupart des traditions et en particulier dans le jeu du *shakuhachi* japonais (voir Castellengo & Fabre, 1994). Entraînés à éliminer les turbulences au profit d'une « note » pure, les flûtistes ont dû adapter leur technique et apprivoiser les oreilles du public aux nouvelles sonorités. Toutefois, les chocs les plus importants sont ceux qui affectent radicalement le mode habituel de production du son. Le jeu aux lèvres à la manière d'une trompette, les sons multiphoniques ou les sons percutés (bruits de clés, tongue-ram) sont autant d'effets sonores surprenants au premier abord mais qui font maintenant partie de l'écriture de l'« instrument flûte » (voir aussi l'ouvrage de Bartolozzi, 1967).

Parmi les compositeurs qui intègrent les bruits d'attaque des instruments dans leur écriture, Helmut Lachenmann a poussé la recherche au point que ses pièces sont qualifiées de « musique concrète instrumentale ».

Ce « démontage » des instruments fait apparaître de nouvelles familles, de nouveaux claviers. Entre un col legno *strident avec beaucoup de pression d'archet et un papier d'emballage vivement déchiré, entre le frottement de deux petits blocs de polystyrène et un souffle brusque dans la clarinette, entre un coup dans l'intérieur du piano et une cymbale raclée, apparaissent des analogies qui dessinent une nouvelle échelle des sonorités, avec laquelle on peut former aussi bien des accords ou des mélodies, comme on le faisait autrefois avec l'échelle chromatique.*

Kaltenecker, M., 1993, *Notice du disque Helmut Lachenmann*.

2.6. Le timbre identitaire et la musique : cohérence et incohérence du timbre

Une certaine cohérence existe entre mode de jeu instrumental et écriture musicale. Nous avons écouté avec plaisir et étonnement la diversité des sons qu'un flûtiste peut jouer dans un exemple musical informel, mais il est vrai que l'écoute du son suivant est déconcertante...

Son 7.31 (7")

Mélodie de timbre « arlequin »

Pourquoi cette mélodie produit-elle un effet surprenant voire cocasse ?

2.6.1. Cocasserie d'un timbre arlequin

Cette fois, l'effet ne provient pas des sons, qui sont ceux du jeu ordinaire, mais de l'incohérence « timbrale » de leur succession dans le déroulement d'une mélodie franchement tonale. Nous avons l'habitude d'associer continuité mélodique et continuité instrumentale. Au sonagramme, le contraste entre l'exemple sonore **7.31** et la même mélodie jouée entièrement à la flûte traversière est frappant.

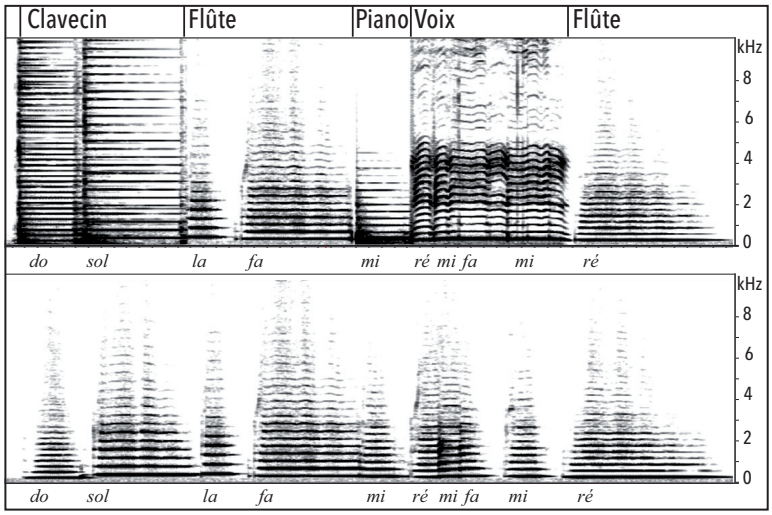


Figure 7.36 Timbre et mélodie. Comparaison de l'analyse sonagraphique de la même mélodie réalisée avec les sons d'instruments différents (en haut) et jouée avec un seul instrument, la flûte traversière (en bas).

L'analyse spectrographique de la mélodie « arlequin » (figure 7.36) produit une image hétérogène qui rend compte des discontinuités instrumentales, en particulier lors du passage du clavecin (*do-sol*) à la flûte (*la-fa*) puis au piano (*mi*), à la voix (*ré-mi-fa-mi*) et à la flûte (*ré*). Le nombre de composantes, la durée d'établissement des sons, les variations expressives (vibrato) changent de façon abrupte. Bien sûr, l'effet de surprise ne vaut que pour la première audition.

Par opposition, la mélodie jouée à la flûte traversière offre une grande cohérence. Confiant dans la stabilité causale, l'auditeur est libre de s'adonner à l'écoute de la musique et du jeu de l'interprète.

2.6.2. Une mélodie et différents instruments

Cohérence du timbre identitaire

Les instruments ayant des caractéristiques très spécifiques sont autant de voix, de « personnages » sonores, reconnaissables aussi par la mélodie et le rythme qui leur sont associés, comme le fit à merveille Prokoviev avec *Pierre et le Loup*. Il est plus intéressant pour notre propos d'écouter ces instruments jouant la même mélodie, ce qui permet de concentrer toute notre attention sur les qualités sonores (Son **7.32**, Son **7.33**). Ce petit exercice permet aussi de comparer d'un seul coup d'œil des instruments différents jouant dans la même tessiture et d'associer l'analyse à une écoute moins ingrate que celle des gammes chromatiques. On trouvera en vis-à-vis sur les figures 7.37 et 7.38 l'analyse de trois instruments à cordes libres et celle de trois instruments à vent.

Les images révèlent d'emblée les différences sonores globales entre ces deux groupes d'instruments. Les trois instruments à cordes libres ont en commun la marque de l'impulsion initiale visible à la fois sur la forme d'onde temporelle et sur l'analyse spectrographique. Ils se différencient principalement par le contenu spectral et la durée de décroissance de la note. La diversité des articulations entre sons successifs – en particulier le legato – se traduit par de subtiles différences dans le contrôle de l'arrêt du son signalé sur la figure par une marque triangulaire rouge.

Le son des instruments entretenus est en permanence sous le contrôle de l'interprète. La durée, les formes d'attaque, le contenu spectral et l'installation du vibrato (flûte, trombone) varient continuellement dans le détail. Il est intéressant de comparer à cet égard, sur chacun des six instruments, la 5e et la 9e note qui est un mi_3. Cependant, à l'échelle de la mélodie (8 s), il se dégage, pour chacun d'eux, une forme globale spectrotemporelle caractéristique.

Son 7.32a (7″)

Clavecin
M. C.

Son 7.32b (8″)

Guitare
E. Pélissier

Son 7.32c (8″)

Piano
M. C.

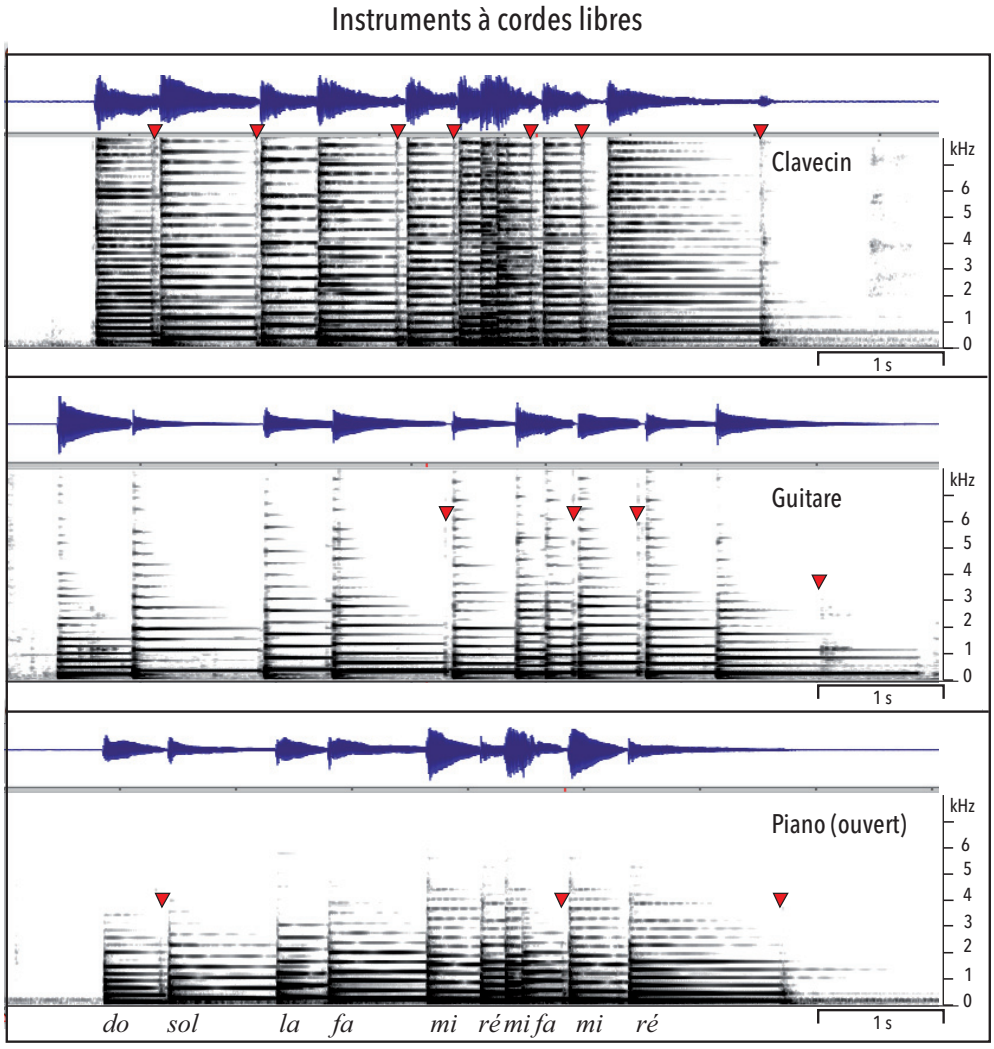
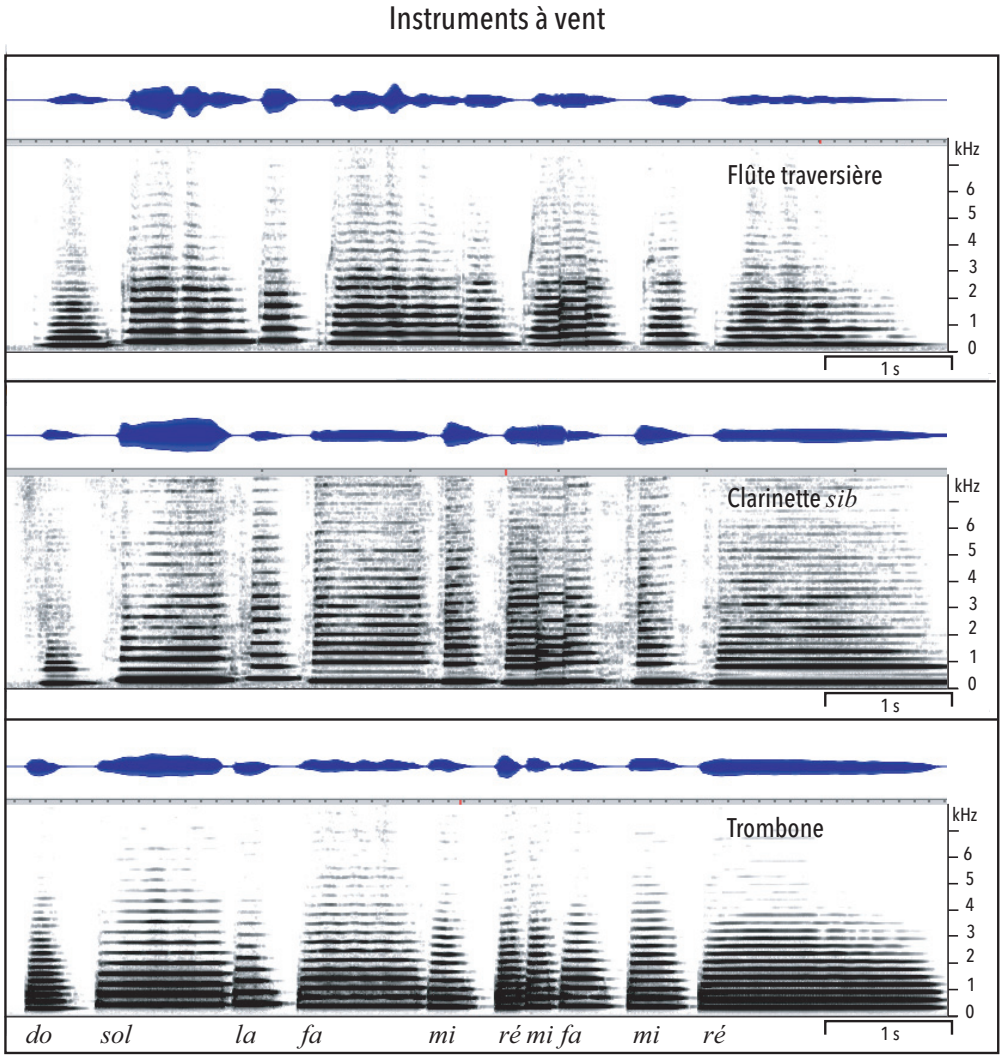


Figure 7.37 Même mélodie jouée sur trois instruments à cordes libres : clavecin, guitare et piano. Les triangles rouges indiquent l'action de l'étouffoir. Forme d'onde temporelle (bleue) et analyse spectrographique. Note de départ do_3. Échelle de fréquence 8 kHz ; temps environ 6 s ; FFT de 2 048 pts.

Il faut préciser ici que le piano et le clavecin, pourvus d'une mécanique de jeu qui limite l'action de l'interprète pour modifier le son, offrent un autre champ de variations qui est celui de la polyphonie. L'étude des variations spectrales, et donc de timbre, qui résultent du mélange de plusieurs sons, est d'une très grande complexité, car la combinatoire est immense. Il devient très difficile d'analyser ce qui relève de l'écoute du « timbre » global d'un agrégat sonore ou de la fonction harmonique de l'accord proprement dit, sans compter l'incidence des interactions qui naissent des rencontres entre fréquences voisines. Les battements et la rugosité peuvent aussi devenir très prégnants temporellement (Pressnitzer, 1998).



Son 7.33a (7″)
Flûte traversière
B. Fabre

Son 7.33b (8″)
Clarinette sib
J.-M. Dussert

Son 7.33c (8″)
Trombone à coulisse
B. Sluchin

Figure 7.38 Même mélodie jouée sur trois instruments à vent. Forme d'onde temporelle (bleue) et analyse spectrographique. Note de départ de la mélodie : do_3. Échelle de fréquence 8 kHz ; temps environ 6 s ; FFT de 2 048 pts.

2.6.3. Le cas particulier de la voix humaine

La voix humaine, qui figure aussi dans les traités d'orchestration, tient une place à part. Les sons vocaux nous sont très familiers et sont immédiatement reconnaissables (voir chapitres 4 et 9), notamment par leur structure acoustique (présence de formants variables et instabilité de la fréquence fondamentale) et pour la forte prégnance de ce signal dans les relations humaines.

Son 7.34 (14″)
Voix chantée (femme). Mélodie sur/a/ puis avec texte.

Le Son **7.34** permet de comparer la même mélodie en chant vocalisé purement mélodique et en chant articulé. Dans la vocalisation à voyelle constante, la voix se comporte comme un instrument possédant des formants fixes (figure 7.39). L'articulation d'un texte produit des modifications acoustiques extrêmement complexes (bruits, variations formantiques, interruptions dues aux consonnes sourdes,

« blancs » spectraux) et suscite des modes d'écoute très différents que chacun peut choisir de privilégier : écoute de la parole et de son interprétation sémantique, écoute de la mélodie, de la voix de l'interprète, de son identité, et de bien d'autres aspects propres à la voix humaine.

Son 7.34 (14")

Voix chantée (femme). Mélodie sur/a/ puis avec texte. M. Garnier

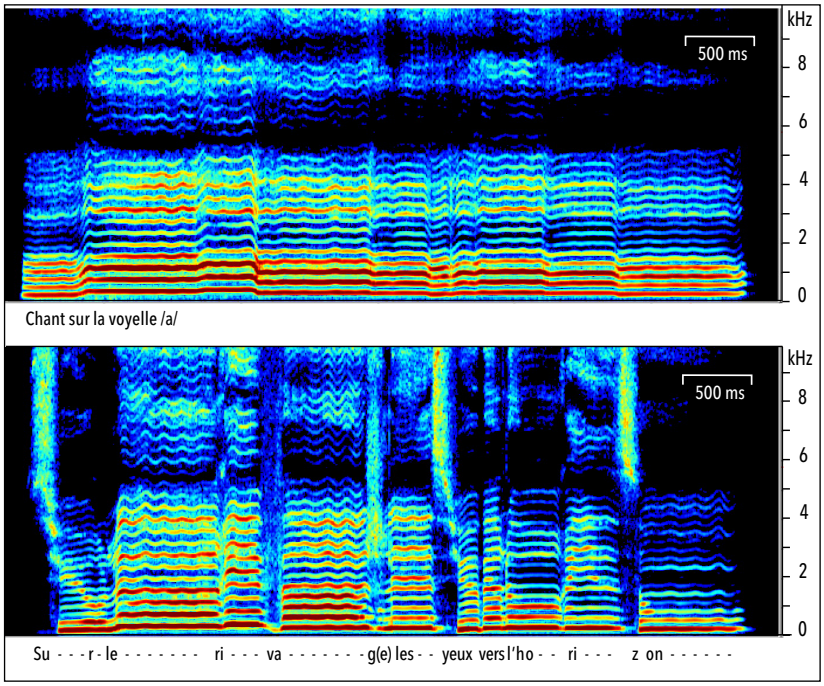


Figure 7.39 Analyses comparées de la mélodie des figures précédentes chantée sur la voyelle « a » tenue, puis chantée avec un texte. Voix féminine.

2.7. Un bilan : typologie instrumentale et timbre causal

Quelles sont les caractéristiques acoustiques pertinentes de la reconnaissance des sons instrumentaux ? Il n'y a pas de réponse générale à cette question. À la diversité des systèmes de production sonore correspondent aussi plusieurs modalités d'appréhension du timbre causal. Au terme de notre exploration acoustique et perceptive, nous avons retrouvé les classes adoptées par les musiciens dans les traités d'orchestration et nous constatons qu'elles s'organisent sur la base des deux critères les plus importants pour la reconnaissance des sources : le type d'excitation et la production (ou non) d'une note plus ou moins précise.

On aboutit donc à quatre catégories principales (la voix humaine étant considérée à part, se référer au chapitre 9) :

1/ les sons de percussion à large spectre, sans hauteur déterminée, dus à une impulsion unique ou répétée,
2/ les sons de percussion à hauteurs multiples plus ou moins bien définies (claviers),
3/ les sons de percussion de hauteur univoque (cordes libres),
4/ les sons entretenus (cordes frottées et vents).

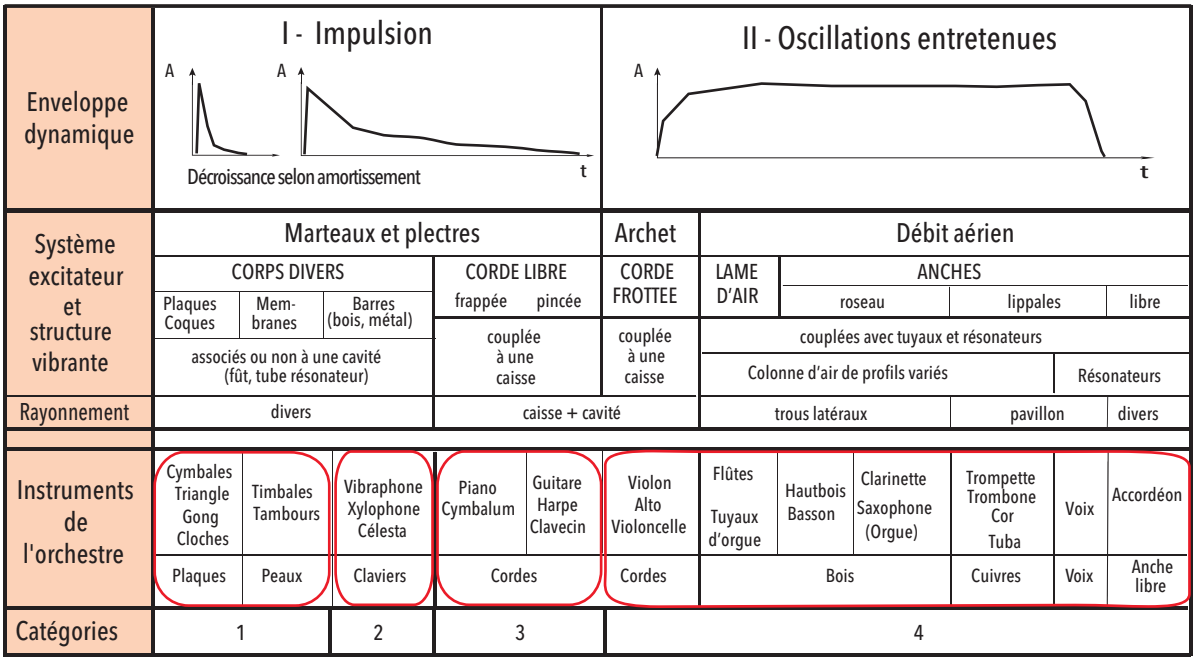


Figure 7.40 Tableau général de l'organisation des principales sources instrumentales de l'orchestre sur la base des principes de production du son. Les quatre groupements de la partie inférieure auxquels nous avons abouti au cours de l'étude du timbre causal répondent aux deux critères pertinents pour la reconnaissance perceptive : la distinction entre son percuté (1, 2, 3) et son entretenu (4) et celle entre les sons de hauteur définie (2) et les sons de hauteur précise (3 et 4).

Ces distinctions tiennent compte du fait que nous avons la possibilité de discriminer et de coder les signaux sonores de deux façons différentes : sur la base de la répartition de l'énergie dans le spectre et sur celle de la détection temporelle des périodicités. Celles-ci conditionnent l'écoute et, ainsi, la diversité des caractérisations du son qui constituent le timbre. Dans les groupes 3 et 4, la « note » produite focalise l'attention auditive, d'autant que la facture instrumentale et les techniques de jeu contribuent à minimiser les autres aspects de la production sonore, considérés comme des phénomènes indésirables. Les modalités spectrales et temporelles de la perception sont alors dévolues à la caractérisation du « timbre causal de la note ». Ce sont, d'une part, le nombre d'harmoniques et la répartition de l'énergie sur ceux-ci et, d'autre part, la façon dont s'établit, se stabilise ou oscille la fréquence de la note.

C'est dans le groupe des instruments à sons entretenus que l'on trouve la plus grande variété de réalisations instrumentales combinant hauteurs et timbres, et que l'interprète a la plus grande marge d'action pendant la durée de production du son. L'expérience de reconnaissance réalisée avec 60 exemplaires de deux notes isolées différentes du groupe 4, puisés dans plusieurs bases de données, a montré le caractère hiérarchisé de la catégorisation opérée par les auditeurs (voir § 2.3.2). Les moins experts d'entre eux discriminent aisément cinq classes ; les plus avertis vont jusqu'à neuf. Le dépouillement des analyses acoustiques et des expériences d'écoute montre une fois de plus la prééminence de la dimension temporelle dans les stratégies de discrimination des sources : non la durée en tant que telle, mais la façon dont varient le spectre et l'amplitude pendant le déroulement temporel du son, ce que nous appelons *la forme temporo-spectrale du son*. En d'autres termes, la durée du transitoire initial d'établissement de

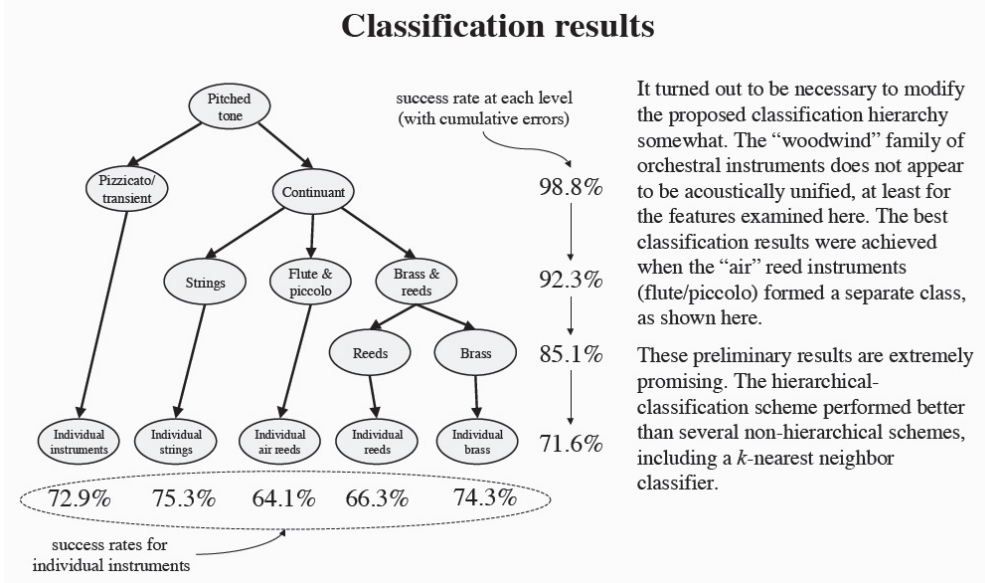
la note varie pour un instrument donné selon l'interprète, mais la succession temporelle des événements qui s'y produisent est caractéristique de l'action mécanique qui est à l'origine du son. Dans cette perspective, les critères de reconnaissance ne peuvent se réduire à une somme de paramètres, même très pertinents. C'est la logique temporelle à laquelle est confronté l'auditeur dans le temps réel de l'écoute qu'il faut débusquer, logique d'enchaînement des événements sonores à partir desquels il repère une forme spectrotemporelle particulière au sein de la catégorie des sons de hauteur définie. Il est plus difficile de définir les critères relevant du contenu spectral de la note proprement dite en raison des nombreuses variables en présence. Les principales sont la tessiture, le rayonnement sonore et le jeu de l'interprète – dont c'est précisément la fonction ! Certaines sources produisent des sons assez similaires sur toute la tessiture : les anches libres, les instruments à embouchure, le violon et, dans une certaine mesure, la flûte traversière. Les instruments à anche et à trous latéraux montrent des variations spectrales par zones (les registres) et d'une note à l'autre, selon la répartition des trous ouverts et fermés.

Qu'est-ce, alors, qu'un instrument de musique ? Schématiquement, c'est une source sonore qui permet de produire des sons ayant en commun la même enveloppe temporo-spectrale de production du son, et en particulier de la note pour les instruments de hauteur. C'est aussi l'ensemble des caractéristiques de comportement qui affectent les sons lorsque sont modifiés l'intensité, la tessiture ou les paramètres de l'excitation, c'est-à-dire lorsqu'on ne se contente plus seulement d'émettre un son, mais que l'on « joue » de l'instrument. Les instruments d'une même famille (violon, alto, violoncelle et contrebasse) partagent le même principe de production et ont donc des enveloppes spectrotemporelles similaires, à quelques facteurs dimensionnels près, puisque chacun d'eux est adapté à une zone donnée de la tessiture.

La reconnaissance du son des instruments de musique intéresse différentes communautés de chercheurs, principalement ceux qui tentent de reproduire le son des instruments par synthèse mais aussi ceux qui s'efforcent aujourd'hui de réaliser des programmes de reconnaissance automatique des instruments par ordinateur. Compte tenu d'une longue tradition d'analyse mathématique du signal sonore, la démarche de ces chercheurs s'inscrit dans le cadre d'une description du signal sonore par paramètres temporels et spectraux, qui sont principalement la durée des transitoires et la description exhaustive du spectre.

Reconnaissance automatique des instruments de musique

Ces dernières années, les efforts des chercheurs ont abouti à la conception de programmes qui traitent automatiquement l'ensemble des sons d'une base de données. La puissance des machines permet d'appliquer plus de cent paramètres. Leur réduction et leur validation sont opérées au cours d'une phase initiale d'apprentissage d'un corpus de sons d'instruments connus. Au cours de la deuxième étape, les algorithmes sont appliqués à un vaste ensemble de sons à identifier. Si les paramètres de la psychophysique de l'audition sont pris en compte, aucune hypothèse cognitive ne guide pourtant ces recherches. La classification automatique fonctionne assez bien pour des sons isolés et, de même que pour les auditeurs humains, le taux de réussite dépend de la distance entre les sons. L'objectif global étant la reconnaissance de la musique, les travaux portent principalement sur les instruments produisant des notes. Tous les auteurs retrouvent la catégorisation initiale du premier niveau (impulsion, entretien) ainsi que les principales catégories du deuxième niveau (cordes libres, cordes frottées, flûtes et généralement aussi le groupe des cuivres). Des divergences apparaissent pour les instruments à anches qui sont différemment regroupés selon les auteurs. Les instruments individuels des niveaux les plus experts ne sont généralement pas reconnus à partir de sons isolés.



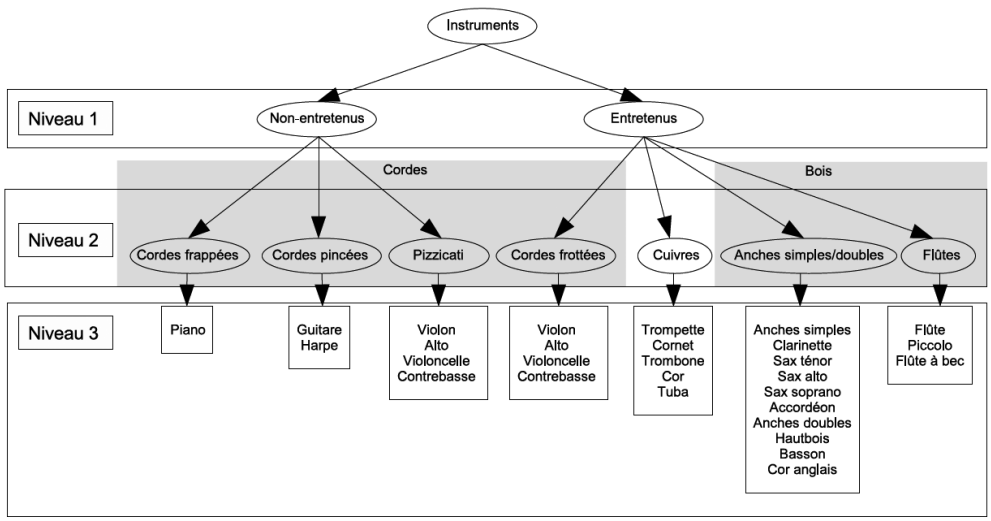


Figure 7.41 Classification automatique des instruments de musique obtenue sur la base d'un apprentissage multicritères. *En haut* : Les résultats de K. Martin (1998) qui a mené aussi des tests de reconnaissance avec des auditeurs. Les catégories obtenues sont similaires aux nôtres, en particulier celle du 2e niveau : cordes (*strings*) ; flûtes (*flutes & piccolo*) ; cuivres et anches (*brass and reeds*). *En bas* : Taxonomie hiérarchique utilisée par Peeters pour la reconnaissance des instruments à partir de notes de musique (2003).

En haut : Martin, K. D., 1998, *Toward automatic sound source recognition : identifying musical instruments*. Poster presented at the NATO Advanced Study Institute on Computational Hearing. En bas : Peeters, G. 2003, cité par Essid, S., 2005, *Classification automatique des signaux audio-fréquences : reconnaissance des instruments de musique*. Thèse UPMC, Paris. (p. 154)

Compte tenu de la démarche adoptée, qui est fondée sur des critères précis permettant de regrouper les sons pour former les classes, il n'est pas étonnant que la prise en compte d'une hiérachisation descendante des sources instrumentales, telle qu'elle se manifeste lors des tests de reconnaissance avec des auditeurs humains, n'apporte aucun gain à l'opération. Elle n'est efficace que pour une stratégie du traitement de la variabilité des formes en temps réel comme l'est la catégorisation perceptive.

3. Le timbre et la catégorisation perceptive

> *Pour construire un timbre de flûte, il faut simuler tous les indices que l'oreille recherche pour confirmer le « fait » que quelqu'un est bien en train de souffler dans une flûte : le souffle, la micromélodie imposée au vibrato par la nervosité de l'instrumentiste, le bruit des lèvres à l'attaque, et ainsi de suite. Lorsque le « fait » est confirmé, l'image mentale d'une flûte naît et rassemble ensuite les multiples sons de timbre très disparates sous l'appellation de « flûte ». On peut considérer que la perception de ces sons disparates comme appartenant à un même instrument est une illusion (car il y a une différence plus grande entre les graves et les aigus d'une flûte qu'entre certaines notes de flûte et de hautbois), un de ces tours que notre esprit épris d'ordre nous joue pour nous rassurer face à un flot chaotique.*
>
> Harvey, J., 1991, p. 456.

Ce tour de passe-passe de notre esprit qui rassemble ensuite les multiples sons de timbre très disparates sous l'appellation de « flûte », et qui fascine Jonathan Harvey n'est pas une illusion : c'est le processus fondamental de la catégorisation perceptive (voir chapitre 4, § 4).

3.1. Les catégories instrumentales du timbre identitaire

3.1.1. L'organisation hiérarchique et les prototypes

Catégoriser, c'est donner du sens, c'est regrouper des éléments apparemment différents sur la base d'un certain nombre de caractéristiques qui leur sont communes et dont la plus grande partie ne se rencontre pas dans les catégories voisines. Ce double mouvement d'agrégation et d'opposition permet une bonne discrimination entre catégories, bien que les frontières entre celles-ci ne soient ni étanches ni fixes dans le temps : elles s'ajustent au gré de l'expérience et selon l'organisation hiérarchique des catégories. Le processus de catégorisation par lequel nous structurons un nombre considérable de sons différents, qui ont en commun des régularités repérables, permet de réduire la variabilité inhérente au monde réel. L'élaboration par chaque auditeur d'un prototype représentatif de chaque catégorie instrumentale réalise ainsi une « économie » cognitive considérable en donnant accès rapidement aux critères essentiels à la reconnaissance ou à la qualification d'éléments nouveaux. Un prototype cognitif sonore renferme à la fois la forme spectro-temporelle caractéristique d'une source acoustique donnée et la façon dont cette forme se transforme en fonction des variables du jeu. Il s'affine, s'enrichit à chaque nouvelle écoute, et ses aspects qualitatifs s'ajustent continuellement au cours des échanges avec les autres musiciens. Comme le dit si justement J. Harvey, il faut bien se mettre d'accord dans une communauté culturelle donnée, sur ce que l'on met sous l'appellation de « flûte », tout en intégrant la diversité des qualités qui lui sont associées.

Niveau de base et niveaux sub-ordonnés

Une catégorisation perceptive s'organise de part et d'autre d'un « niveau de base »[42], celui pour lequel les regroupements par similarité sont les plus communément partagés par les membres d'une communauté et dont le prototype représentatif de chaque catégorie correspond aux désignations verbales les plus courantes (voir figure 7.42). Pour les instruments de musique, on y trouve : le piano, la guitare, le violon, etc. (pour plus d'informations voir chapitre 4, § 4.2).

42. « La notion de niveau de base ne correspond pas toujours à la classification musicologique : pour un grand nombre d'auditeurs (et il semble que ce soit d'autant plus important que ceux-ci sont néophytes en musique), « trompette » sert à identifier l'ensemble des cuivres, pour d'autres, l'ensemble des sons obtenus en soufflant (équivalent des vents), alors que certains utilisent bien ce terme pour identifier uniquement le son TPT, c'est-à-dire au niveau de l'instrument de musique précis. » (Anne Faure, 2000, p. 356)

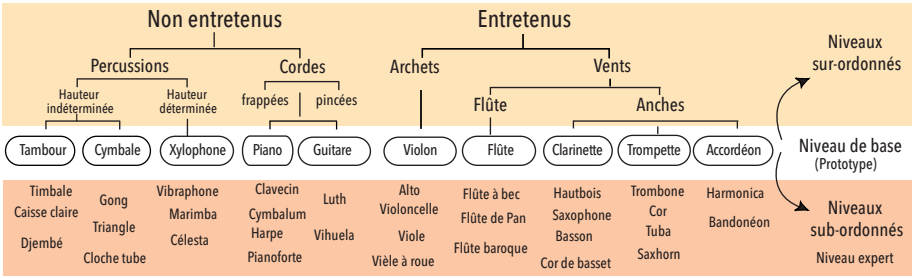


Figure 7.42 Schéma d'une catégorisation des instruments de musiques usuels. Au niveau de base, certains instruments jouent le rôle de prototype pour les instruments proches mal différenciés.

Lorsqu'une catégorie s'enrichit, se complexifie, il devient nécessaire de l'organiser en créant des sous-catégories. Peu à peu des niveaux sub-ordonnés se constituent, en nombre variable selon les auditeurs. Citons, par exemple, les catégories d'instruments différents mais de sonorités voisines, comme les cors, les saxhorns et les bugles, ou encore les instruments répondant à des esthétiques sonores relevant d'époques différentes (pianoforte, viole) ou de cultures différentes (vièle à roue, musette, djembé). D'autres catégories de niveaux sub-ordonnés rendent compte des variations de sonorité d'instruments d'une même famille, donc de même facture, mais qui diffèrent par leurs dimensions : par exemple les instruments du quatuor à cordes ou ceux de la famille des saxophones. Prenons quelques exemples concrets en commençant par la comparaison d'instruments de la même famille : les flûtes traversières et les cordes frottées.

3.1.2. Les niveaux de catégorisation et l'expertise : les familles d'instruments

Les instruments à vent et à cordes frottées, dans lesquels le changement de hauteur sonore est dû à des transformations du corps sonore, ont une étendue limitée (trois à quatre octaves). Il en existe de différentes tailles, chaque modèle ayant ses caractéristiques sonores.

Les flûtes traversières. Dans l'exemple **7.35**, on entend successivement quatre sons isolés, de même fréquence, prélevés dans une prestation musicale jouée par le même flûtiste. Étant donné les possibilités de variations de la qualité sonore dont dispose un flûtiste, il est assez difficile de savoir, à l'écoute d'une seule note, si les variations perçues sont dues à l'interprète ou si celui-ci joue des instruments différents.

Son 7.35 (8")

Quatre sons de flûte : est-ce le même instrument ?

Pour les auditeurs experts que sont les flûtistes, il ne fait aucun doute qu'il s'agit d'instruments différents, comme le montre l'écoute de la phrase musicale complète (Son **7.36**). L'attaque, les transitions entre les notes et la coloration du souffle permettent de discriminer les différences de comportement des instruments, de même que certains indices comme l'attaque sur l'octave inférieure qui indique un fonctionnement en régime 2 (figure 7.43, flûte 3).

Son 7.36 (24")

Mélodie jouée successivement avec quatre flûtes de tessitures différentes

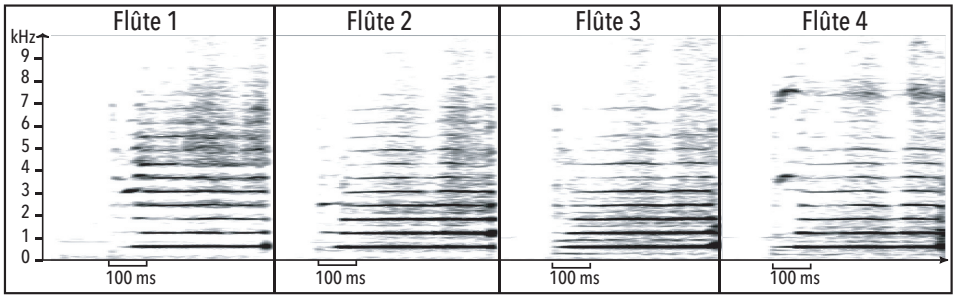


Figure 7.43 Analyse des quatre *ré*$_4$ de l'exemple sonore 7.35. Comparez les zones de souffle, le contenu harmonique, le son de bouche du transitoire d'attaque. Notez aussi l'attaque de la flûte 3, commençant à l'octave inférieure (analyse FFT 2 048 pts ; 10 kHz).

Son 7.37 (14″)

Note finale de chacune des quatre mélodies du Son 7.36

Pour se convaincre de l'importance des indices spectrotemporels du transitoire d'attaque, on écoutera le Son **7.37** qui est un montage en désordre de la dernière note tenue[43]. La durée d'attaque de chaque note est similaire (environ 200 ms).

Les instruments du quatuor. La séquence sonore suivante permet d'écouter un exemple similaire, c'est-à-dire une mélodie de hauteur donnée jouée sur trois instruments de dimensions différentes de la famille « violon ». La mélodie est interprétée sur un violon, un alto et un violoncelle par leurs instrumentistes respectifs (Sons **7.38a**, **b**, **c**).

Une même mélodie jouée aux instruments du quatuor

Son 7.38a (13″)

Violon

Son 7.38b (12″)

Alto

Son 7.38c (12″)

Violoncelle

Par comparaison aux instruments à trous latéraux, les variations de sonorité dues au fait que les instruments jouent dans des zones différentes de leurs tessitures se traduisent principalement par des variations de « balance spectrale » entre les zones graves et aiguës du spectre. La figure 7.44 montre l'analyse spectrale IDS moyennée sur la durée totale de la phrase musicale (voir le chapitre 2, § 3.5.2). Les trois instruments couvrent à peu près la même largeur totale de spectre, mais la contribution énergétique aux basses fréquences diffère selon les dimensions de l'instrument : elle est plus importante pour le violoncelle que pour le violon.

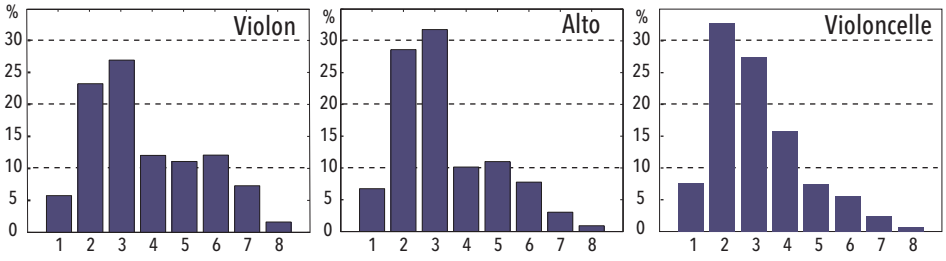


Figure 7.44 Analyses moyennées (IDS) de la même phrase musicale jouée sur trois instruments du quatuor. En ordonnée : énergie en pourcentage de l'énergie totale. En abscisse : n^{os} des bandes de fréquence. La contribution énergétique des basses fréquences croît avec la taille de l'instrument, particulièrement pour le violoncelle (voir la bande 2, 200-400 Hz). Le caractère « brillant » du violon tient à la contribution de la bande 6 (1 800-3 000 Hz) ciblant la zone sensible de l'oreille. (Séance de prise de son avec Albert Laracine et les élèves de la Fémis)

43. Les flûtistes auront peut-être reconnu dans l'ordre : flûte basse, flûte en *ut*, flûte alto et piccolo.

Les quatre flûtes de notre exemple et les trois instruments du quatuor sont des instruments homothétiques adaptés à des tessitures différentes. Ils relèvent d'une même catégorie identitaire du niveau de base, respectivement la « flûte » et les « cordes frottées de type violon ». Les noms que l'on donne au violon, à l'alto, au violoncelle désignent effectivement des instruments distincts, mais ils correspondent à des variantes sonores assez fines lorsqu'on les entend dans la même tessiture. Il en est de même pour la trompette et le trombone[44] de la classe des « cuivres » ou encore pour certains jeux d'orgue comme le prestant et la doublette qui sont des répliques, à une et deux octaves de distance, de la montre. Clark et son équipe (Clark, Robertson & Luce, 1964) ont montré par des tests d'identification que la plupart des auditeurs confondaient les instruments d'une même famille lorsqu'ils étaient joués à la même hauteur. Pourtant, des musiciens experts, en l'occurrence ceux qui pratiquent ces instruments, peuvent les discriminer sans peine[45].

Que des instruments soient tantôt perçus comme des variations de qualité d'un type d'instrument donné ou comme des exemplaires de catégories différentes n'est pas contradictoire avec le fait que reconnaître et qualifier sont des perceptions antagonistes ou qui, du moins, mettent en jeu des stratégies d'écoute différentes. Les mêmes signaux font sens différemment pour des auditeurs musiciens peu compétents en instrumentation et pour des auditeurs experts capables d'effectuer de nombreuses discriminations entre les instruments. Les premiers ont pour référence un prototype global « instrument à cordes frottées » sur la base duquel ils apprécient des qualités différentes. Les seconds, qui ont accumulé un grand nombre d'indices au cours de leur pratique ont constitué des sous-catégories identitaires, reliées hiérarchiquement à la catégorie principale, qui possèdent également leurs prototypes de qualité sonore : ce sont respectivement celles du violon, du violoncelle ou celle de l'alto (voir § 3.1.1).

3.1.3. Un exemple de prototype qualitatif du timbre : le son d'une guitare baroque

La mémoire des références culturelles, ainsi que l'apprentissage instrumental, jouent donc un rôle primordial dans l'élaboration des prototypes de référence. Lorsque la pratique d'un instrument ancien s'est perdue ou lorsqu'il n'existe pas de référence en mémoire, l'appréciation se fait nécessairement sur la base des références sonores du moment[46]. Le regain d'intérêt pour les musiques anciennes des XVIe et XVIIe siècles a permis d'en prendre la mesure. Aujourd'hui, des instruments de facture ancienne côtoient des instruments modernes qui portent le même nom : flûte, guitare, violon, mais dont l'esthétique sonore est parfois très différente.

Son 7.39 (21″)

Guitare baroque, enregistrement A

Dans les extraits musicaux suivants (Son **7.39** et Son **7.40**), nous entendons une guitare d'esthétique baroque, construite par Charles Besnainou, qui a fait l'objet de deux enregistrements (A et B) par le même interprète. Bien qu'il s'agisse de musiques différentes, nous avons peine à croire qu'il s'agit bien du même instrument.

Son 7.40 (22″)

Guitare baroque, enregistrement B

La qualité professionnelle des deux ingénieurs du son n'étant pas en cause, que s'est-il passé ? La seule explication possible est que le preneur de son de l'enregistrement B s'est efforcé d'« arranger » le son afin de le rapprocher au mieux de celui de la guitare moderne qu'il avait en tête. Il a donc atténué fortement les fréquences aiguës et accentué la partie grave du spectre pour répondre à l'idéal actuel.

44. Le trombone est un instrument cylindrique à embouchure ; c'est la basse de la trompette du point de vue de la sonorité.
45. Dans les test d'écoute d'identifications d'instruments de musique sur une seule note, plusieurs auteurs comptabilisent de la même façon les confusions des auditeurs entre flûte et clarinette, ou entre violon et alto.
46. Il est étonnant d'entendre aujourd'hui les premiers clavecins Pleyel construits au début du XXe siècle, sortes de « pianos » à cordes pincées, puissants, avec beaucoup de graves. Voir le site du musée de la Musique de Paris et pour écouter : http://www.mim.be/fr/clavecin-pleyel ?from_i_m=1.

Des problèmes similaires se manifestent lorsque de nouveaux instruments apparaissent ou lorsque des modifications de facture, destinées à faciliter le jeu ou à perfectionner la justesse, ont des retentissements soudains et trop importants sur la sonorité admise habituellement. Les changements apportés à la flûte traversière par Th. Boehm ont rencontré beaucoup de résistance pendant au moins deux générations (Castellengo M. & Forest L, 1999). Quand au basson Boehm, ce fut un échec (Kergomard J., 1976).

Qu'il s'agisse de reconnaître une source ou d'être capable d'en caractériser les qualités, la sûreté et la finesse de performance de l'auditeur seront d'autant plus remarquables que l'expérience d'un grand nombre d'écoutes et de leur mémorisation auront pu être développées.

3.1.4. La catégorisation hiérarchique et les stratégies d'analyse

La source ou la sonorité qu'il s'agit de caractériser peut se définir à différents niveaux de catégorisation selon les contextes d'écoute (figure 7.45). S'agit-il de comparer le timbre de la flûte par rapport à celui de la clarinette ? Celui de la flûte traversière par rapport à celui de la flûte à bec ? Le timbre de la flûte Boehm par rapport à celui de la flûte traversière baroque ? Ou les timbres de deux notes différentes comme le *sol#* et le *la* d'une même flûte ?

On conçoit qu'il faudra adapter différemment les stratégies d'analyse acoustique selon la question posée. On remarquera, par ailleurs, que plus les classes de sources sont voisines, plus les variables externes (conditions de prise de son, rôle du musicien) devront être maîtrisées pour mettre en évidence celles qui sont dues seulement à l'instrument.

Classes de comparaison	Niveau	Critères
Instrument à vent flûte — clarinette	Niveau de base	*contenu spectral*
Flûte à bec — traversière	Niveau sub-ordonné 1	*transitoire* *composition spectrale* *stabilité du son* *homogénéité de tessiture*
Flûte traversière baroque — syst. Boehm	Niveau sub-ordonné 2	*richesse spectrale* *puissance sonore* *homogénéité chromatique* *justesse*
Flûte traversière syst. Boehm *sol#* — *la*	Discrimination locale	*égalité sonore* *égalité d'intensité*

Figure 7.45 Lorsque nous caractérisons les différences entre deux sons, nous adaptons les critères acoustiques en fonction de leur « distance perceptive ».

3.2. Caractériser le timbre qualitatif : les mots du timbre

3.2.1. Les traités de musique

Marin Mersenne (1636)

Avant qu'il soit possible d'enregistrer les sons pour analyser la sonorité des instruments, les musiciens avaient pour seule ressource de décrire l'impression produite par le son au moyen du langage. Encore faut-il définir de quel « son » il s'agit. Dans *l'Harmonie universelle*, Marin Mersenne formule ainsi la question : « À savoir quel est le plus agréable son de tous les instruments de musique ». Il précise que la « douceur et l'excellence du son considéré tout seul est le sujet unique de ce discours » mais s'empresse d'ajouter : « encore qu'il soit comme impossible de bien juger de ce son sans parler de la bonté des instruments, l'un étant tellement joint à l'autre qu'ils n'en peuvent être séparés »[47]. L'appréciation d'un son est donc relative à une catégorie instrumentale donnée. Dans le cours du texte, on relève divers adjectifs (doux, agréable, rude, gai, aigre*)* et certaines expressions (la « douceur morne et sombre de la flûte »[48], opposée à la « gaité naturelle » des anches) qui rendent compte assez clairement de la dimension hédonique – plaisir ou désagrément – du son d'un instrument et de l'imaginaire qui lui est associé.

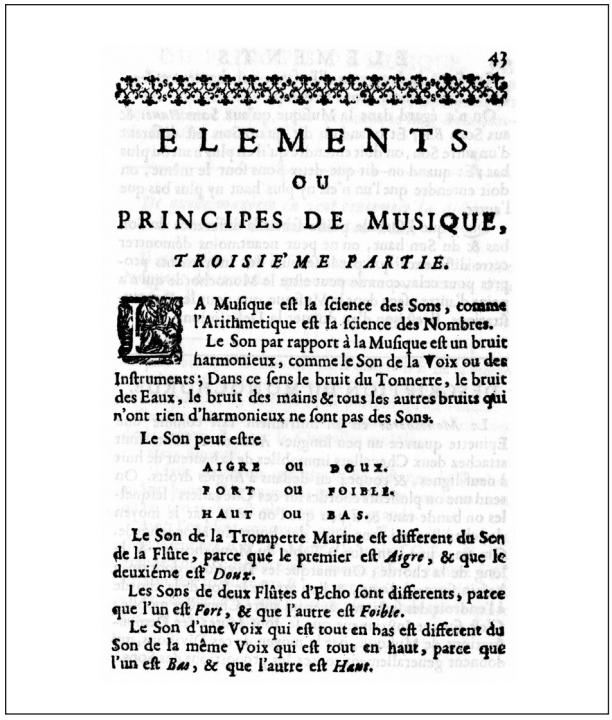
Étienne Loulié (1696)

Avec l'intérêt suscité par l'acoustique, les auteurs vont s'efforcer progressivement de caractériser plus précisément le son en tant que phénomène à part entière. En 1690, É. Loulié écrit « Trois sortes de Son. Son Physique, Son Mathématique, Son de Musique (...) Son de Musique est celuy qu'on peut mesurer par l'Oreille comme le son de la violle, de la flute, de la voix » [49]. Le même auteur écrit en 1696 (figure 7.46) : « Le Son peut estre AIGRE ou DOUX », ce qui, à peu de choses près, est la formulation que l'on retrouve plus tard sous la plume de J.-J. Rousseau (1752) pour le nouveau terme « timbre », à l'article « son » de *l'Encyclopédie* de Diderot et d'Alembert (voir encadré page 289).

Les discours sur les qualités des sons sont particulièrement intéressants et riches dans le domaine de la facture d'orgues puisqu'il s'agit de décrire les défauts individuels des tuyaux et de donner les moyens d'y remédier, tout en se basant sur l'écoute et sur une description verbale des qualités sonores. Dans *l'Art du facteur d'orgues* (1770), Dom Bédos commence ainsi le chapitre X intitulé « Faire parler les tuyaux à bouche » : « C'est ici le chapitre le plus important puisqu'il s'agit du son de l'orgue. » L'auteur n'utilise à aucun moment le terme timbre. Ses commentaires concernent tantôt le « son » d'un tuyau qui est « sec, maigre, grossier, etc. », tantôt le « tuyau » lui-même qui « souffle, crie, parle nettement, est sujet à friser, nasarde, etc. ». L'objectif à atteindre est la belle « harmonie », l'harmonie *véritable* qui dépend du caractère prototypique de chaque jeu : le « moelleux et la douceur » des flûtes ; le caractère « mâle et éclatant » des anches.

Il est manifeste qu'un terme unique comme le « timbre » ne peut rendre compte des différentes appréhensions des qualités du son qui s'inscrivent dans une pratique réelle, lorsque écouter et qualifier les sons font sens dans le cadre d'une activité, pour un projet sonore défini. De surcroît les auteurs s'efforcent de distinguer les deux parties du son, c'est-à-dire le début (attaque) et la partie tenue. Mersenne analyse séparément la « percussion » du son du luth et sa « résonance »[50]. Dom Bédos insiste à plusieurs reprises sur l'alliance de « la promptitude à parler avec la belle harmonie » pour les jeux d'anche, car c'est le facteur d'orgue qui décide pleinement de la sonorité de chaque tuyau, laissant au musicien la possibilité de combiner les jeux et les claviers, selon son goût et selon les musiques.

47. Marin Mersenne, *l'Harmonie universelle*, Livre premier des instruments, Proposition IIII.
48. Il s'agit d'une flûte à 6 trous cylindrique, de type *Renaissance.*
49. Voir Psychoyou Th., 2003, tome I, p. 194.
50. Livre premier des instruments, Proposition IIII.



43

ELEMENTS
OU
PRINCIPES DE MUSIQUE,
TROISIE'ME PARTIE.

LA Muſique eſt la ſcience des Sons, comme l'Arithmetique eſt la ſcience des Nombres.

Le Son par rapport à la Muſique eſt un bruit harmonieux, comme le Son de la Voix ou des Inſtruments; Dans ce ſens le bruit du Tonnerre, le bruit des Eaux, le bruit des mains & tous les autres bruits qui n'ont rien d'harmonieux ne ſont pas des Sons.

Le Son peut eſtre

AIGRE ou DOUX.
FORT ou FOIBLE.
HAUT ou BAS.

Le Son de la Trompette Marine eſt different du Son de la Flûte, parce que le premier eſt *Aigre*, & que le deuxiéme eſt *Doux*.

Les Sons de deux Flûtes d'Echo ſont differents, parce que l'un eſt *Fort*, & que l'autre eſt *Foible*.

Le Son d'une Voix qui eſt tout en bas eſt different du Son de la même Voix qui eſt tout en haut, parce que l'un eſt *Bas*, & que l'autre eſt *Haut*.

Figure 7.46 Extrait de l'ouvrage *Éléments ou principes de musique mis dans un nouvel ordre*.
É. Loulié, 1696, p. 43 [Minkoff].

Traités d'instrumentation

Lorsque l'orchestre – que l'on pourrait qualifier d'hyperinstrument – se développe, les sonorités des différentes familles d'instruments tendent à se stabiliser, voire à se standardiser, au bénéfice d'un travail sur leur combinatoire, extrêmement complexe et demandant beaucoup d'expérience musicale. Les premiers traités de « science musicale du timbre », selon l'expression de G. Kastner (1837)[51], paraissent dès la première moitié du XIXe siècle pour « venir en aide aux jeunes compositeurs ». Peu après, Berlioz introduit son *Traité d'instrumentation et d'orchestration* (1844) en annonçant « l'étude fort négligée jusqu'à présent, de la nature du timbre, du caractère particulier et des facultés expressives de chacun d'eux » (les instruments). Pour faire état de la complexité de l'entreprise nous extrayons très arbitrairement quelques exemples de « description du timbre » empruntées à Ch. Koechlin (1935, vol.1). Ces ouvrages retrouvent aujourd'hui un regain d'intérêt dans le cadre de recherches visant à puiser dans les bases de données numérisées, des sons dont la qualité est définie par quelques adjectifs (Carpentier, 2008).

Charles Koechlin ouvre son *Traité de l'orchestration* par l'étude de la grande flûte. La tessiture de l'instrument est divisée en zones caractérisées par des qualificatifs (figure 7.47). Certains sont aisés à interpréter. Le terme « sifflant » évoque vraisemblablement les bruits d'attaque et le souffle des sons graves ; celui « d'incisif » se rapporte au fait que les sons du suraigu, proches de la zone sensible de l'oreille, émergent facilement de l'orchestre. En revanche, on ne saisit pas bien pourquoi Koechlin place un changement de sonorité au la_3 ni pourquoi il précise dans le texte que le si_4 est « nettement plus clair » que le *la* qui le précède, et qu'il est le

51. Voir Woodward, P. J., *Jean-Georges Kastner's, Traité général d'instrumentation* ; A Translation and Commentary. Denton, Texas. UNT Digital Library. http://digital.library.unt.edu/ark:/67531/metadc4165/ ; consulté en 2010.

premier d'une série de sons « lumineux ». Les flûtistes de l'entourage de Koechlin marquaient vraisemblablement ces différences de qualité sonore qu'il nous est difficile d'imaginer aujourd'hui. Plus loin Koechlin décrit les qualités du basson[52] avec une grande abondance de termes : « Notes du grave pleines et bien timbrées, (...) sons excellents, parfaits, avec un charme extrême, voilé, clair, timbre plein et velouté, notes tendues, harmonieuses, sonorité nostalgique, émouvante » (p. 42).

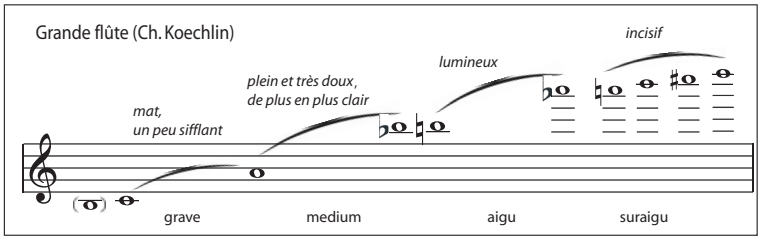


Figure 7.47 Qualifications du son de la flûte traversière selon les zones de sa tessiture.

D'après Koechlin Ch., 1935, vol. 1, page 6.

Un instrument est toujours décrit pour sa sonorité et pour ses qualités propres à servir des caractères, à susciter des émotions. L'entreprise est d'autant plus difficile que les effets changent selon la tessiture, selon l'intensité de jeu, selon les modes de jeu et les articulations. Aussi les premières recherches objectives fondées sur l'analyse acoustique du son ont-elles suscité beaucoup d'espoir.

3.2.2. Helmholtz : le « timbre musical » de la note

La première étude expérimentale du timbre est due à Hermann von Helmholtz, physicien et physiologiste. Partant de la théorie mathématique de Fourier (1822) dont il démontre le fondement physique à l'aide de résonateurs isolant les harmoniques, il établit pour la première fois un lien expérimental entre le fonctionnement physique des systèmes excitateurs des instruments de musique, le contenu harmonique des sons et les qualités habituellement associées aux différents instruments.

> *Le but du présent chapitre sera donc d'analyser les différents sons qui proviennent des instruments de musique, afin de mettre en évidence les différents caractères que l'on observe dans les combinaisons des harmoniques ou notes supérieures, correspondant à des variétés caractéristiques du timbre. Nous serons amenés à un certain nombre de règles générales de ces arrangements de notes supérieures, règles qui répondent dans le langage à différentes manières d'être du timbre*[53] *que l'on désigne par les mots : timbre doux, aigu, éclatant, creux, plein ou riche, sourd, clair, etc.*
>
> Helmholtz (von), H., 1874, p. 97.

L'ouvrage contient plusieurs remarques très intéressantes sur la perception des sons, en particulier sur le rôle de l'attention auditive qui nous permet d'écouter isolément les harmoniques d'un son ou, au contraire, de les appréhender en bloc.

52. En contraste total avec les qualificatifs dépréciatifs de Berlioz sur cet instrument qu'il n'appréciait guère.
53. Dans les textes de l'époque, le terme allemand *klangfarbe* (littéralement « couleur du son »), est traduit en français par « timbre », et en anglais par « sound quality ». L'ambiguïté est totale entre les deux aspects du timbre qui sont l'identité, ou qualité des sons (voir note de la page 290)

Le timbre musical selon Helmholtz

Dans le chapitre 5 de la *Théorie physiologique de la musique* (édition française, 1874), Helmholtz consacre plusieurs pages à la description de traits acoustiques audibles qui interviennent dans le timbre des sons, et conclut qu'il doit limiter son étude à la partie « régulière » des vibrations celle qu'il nomme le « timbre musical ».

Nous ferons abstraction, dans ce chapitre, de toutes les irrégularités du mouvement de l'air, ainsi que des circonstances relatives au commencement et à la fin du son, et nous n'aurons égard qu'à la partie du son véritablement musicale, celle qui correspond à un mouvement d'une durée constante et de période régulière. Nous rechercherons les rapports qui existent entre la constitution du mouvement vibratoire et le timbre du son qui en résulte, et le terme de *timbre musical* ne devra rappeler que les particularités dont nous allons nous occuper.

Figure 7.48 Helmholtz mentionne ici les phénomènes qui contribuent à l'appréciation du timbre mais qu'il ne peut prendre en compte en raison de leur variabilité (Helmholtz, H., 1874, *op. cit.*, page 96).

À l'époque d'Helmholtz, seuls les sons tenus sont analysables, comme ceux de tuyaux d'orgues posés sur un sommier.

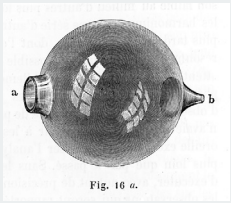


Figure 7.49 Exemple de résonateur sphérique servant à l'analyse des harmoniques à l'oreille.
Helmholtz, H., 1874, *op. cit.*, page 59.

L'analyse des harmoniques se pratique à l'oreille, à l'aide d'une série de diapasons de différentes fréquences, et de résonateurs (figure 7.49). Il faut disposer d'un résonateur par harmonique. L'analyse des six premiers harmoniques d'un do_2 mobilise les résonateurs accordés respectivement sur les fréquences de do_2, do_3, sol_3, do_4, mi_4 et sol_4. Une des extrémités du résonateur, terminée en pointe (b), est introduite dans le conduit de l'oreille et colmatée au pavillon avec un peu de cire. Helmholtz précise qu'il bouche l'autre oreille, également avec de la cire. Il faut répéter l'opération pour l'analyse de chaque harmonique, ce qui est très long.

Un grand nombre de ces remarques a trait aux phénomènes qui se produisent au début et à la fin du son, aux bruits et autres irrégularités temporelles. Cependant, contraint par les limites des moyens expérimentaux dont il dispose, l'auteur ne peut étudier que la partie tenue des sons, ce qu'il nomme « le timbre musical »[54].

Pour définir les « caractères de sons musicaux » (voir encadré page suivante), Helmholtz prend en compte le nombre d'harmoniques, l'intensité relative du fondamental par rapport aux harmonique supérieurs et la prédominance des harmoniques impairs. Il établit des correspondances entre les paramètres physiques issus de ses analyses et les adjectifs censés représenter au mieux ces caractères. Les voyelles, qui « *se distinguent essentiellement des sons de la plupart des autres instruments de musique, en ce que l'intensité des harmoniques qu'elles présentent ne dépend pas du numéro d'ordre, mais de la hauteur absolue* », font l'objet d'une étude séparée.

54. Malgré les précautions qu'il prend pour restreindre ses résultats aux conditions expérimentales dans lesquelles il a travaillé, ceux-ci seront simplifiés et repris de façon caricaturale dans les ouvrages ultérieurs. Aujourd'hui encore, il n'est pas rare de voir que le timbre d'un instrument donné est assimilé au spectre harmonique d'une seule « note », sans qu'il soit fait mention de la dimension temporelle.

Règles générales qui unissent le timbre à la composition du son selon Helmholtz[1]

Note : nous avons mis en caractères gras les termes utilisés par Helmholtz pour caractériser le timbre.

1° Des sons simples, comme ceux des diapasons associés à des tuyaux résonants, ceux des grands tuyaux bouchés de l'orgue, présentent beaucoup de **douceur**, de **charme**, n'ont aucune **dureté**, mais ils manquent **d'énergie**, et sont **sourds** dans les régions graves.

2° Des sons accompagnés d'une série d'harmoniques graves de moyenne intensité, jusqu'au sixième environ, sont **pleins** et d'un bon emploi en musique. Comparés aux sons simples, ils ont quelque chose de plus **riche**, de plus **fourni**, et sont cependant parfaitement **harmonieux** et **doux**, tant que les harmoniques supérieurs font défaut. À cette catégorie appartiennent les sons du piano, des tuyaux ouverts de l'orgue, les sons faibles et doux de la voix humaine et du cor ; ces derniers formant la transition du côté des sons munis d'harmoniques élevés, tandis que les flûtes et les jeux de flûte, avec peu de vent, se rapprochent des sons simples.

3° Quand les sons partiels impairs existent seuls, comme dans les petits tuyaux bouchés de l'orgue, les cordes du piano pincées au milieu et la clarinette, le son prend un caractère **creux** et même **nasillard**, pour un grand nombre d'harmoniques.

5° Si le son fondamental domine, le timbre est **plein** ; il est **vide**, au contraire, si l'intensité du son fondamental ne l'emporte pas suffisamment sur celles des harmoniques. Ainsi le son des grands tuyaux ouverts de l'orgue est plus plein que celui des petits tuyaux de même nature ; le son des cordes est plus plein, lorsqu'elles sont ébranlées par les marteaux du piano, que lorsqu'elles sont frappées avec un morceau de bois ou pincées par les doigts ; le son des tuyaux à anche, associés à des appareils résonnants appropriés, est plus plein que celui des mêmes tuyaux sans caisses résonnantes.

4° Quand les harmoniques supérieurs, à partir du sixième ou du septième, sont très nets, le son devient **aigre** et **dur**. Nous en trouverons plus tard l'explication, dans les dissonances que forment entre eux ces harmoniques supérieurs. Le degré de **mordant** peut varier ; avec une faible intensité, les harmoniques supérieurs ne diminuent pas essentiellement la possibilité de l'emploi musical du son ; ils augmentent, au contraire, le caractère et la puissance d'expression de la musique. Dans cette catégorie figurent, avec une importance particulière, les sons des instruments à archet, puis, la plupart des instruments à anche, le hautbois, le basson, l'harmonium, la voix humaine. Les sons **durs** et **éclatants** des instruments de cuivre sont extraordinairement **pénétrants**, et, par suite, donnent l'impression d'une grande **puissance**, à un plus haut degré que les sons de même hauteur, mais d'un timbre **doux**. Aussi conviennent-ils peu, par eux-mêmes, à la musique de chambre, mais ils sont d'un grand effet à l'orchestre. On verra plus tard de quelle manière les harmoniques supérieurs peuvent rendre le son plus pénétrant.

1. Helmholtz, *op. cit.*, 1874, p. 150

Ce travail, effectué plus de quinze ans avant le premier enregistrement du son, dans lequel l'auteur combine des expériences rigoureuses à de longues et patientes analyses à l'oreille, marque une étape importante dans la connaissance du contenu harmonique des sons. Pour rendre sensible les caractères du son, Helmholtz utilise des adjectifs et cite des exemples d'instruments de musique, mais il lui manque de pouvoir valider ses résultats en les expérimentant avec des auditeurs.

3.3. Les attributs perceptifs du timbre : des sons synthétiques aux sons instrumentaux

Nous utilisons ici le terme « timbre » sans plus de précision, tel que les auteurs l'emploient.

3.3.1. La recherche expérimentale sur le timbre avec les tests d'écoute

De 1940 à nos jours, une quantité considérable de travaux a été effectuée. Les étapes en sont rythmées par les progrès techniques de l'instrumentation acoustique – de l'électronique à l'informatique – ainsi que par le développement des connaissances sur les méthodes de tests et le traitement statistique des résultats. Pour plus d'informations, nous engageons le lecteur curieux à se reporter aux articles de J. Hajda *et al.* (1994), et de S. Donnadieu (2007).

La quasi-totalité des recherches effectuées sur la perception du timbre se fonde sur les paradigmes de recherche de la psychophysique dont l'objectif est de mettre en évidence les « dimensions » perceptives du timbre. Dimensionnel s'oppose ici à catégoriel. Cela signifie que lorsqu'un paramètre physique donné varie de façon graduelle, les changements perceptifs doivent se produire aussi de façon continue et qu'une relation cohérente entre le stimulus et la réponse perceptive peut être formulée. Pour réaliser cet objectif, les chercheurs doivent créer des sons égalisés perceptivement en hauteur, en intensité et en durée (ce qui impose de travailler avec des sons isolés). Ces sons diffèrent entre eux par certains aspects spectraux ou temporels qui sont les paramètres visés. La situation décontextualisée du test en laboratoire ainsi que les procédures expérimentales de calibrage des sons favorisent de fait une écoute analytique et paramétrée. Les auditeurs sont considérés ici comme des récepteurs psychophysiologiques doués de capacités auditives et de seuils de discrimination qu'il s'agit de mesurer. Un nombre suffisant d'auditeurs est nécessaire pour observer des régularités statistiquement significatives dans les réponses.

3.3.2. Les tests avec sons synthétiques paramétrés : la « brillance »

Pour mettre en place des tests d'écoute, il faut pouvoir créer des sons dont on pourra contrôler les paramètres acoustiques. Les premiers chercheurs (Lichte, Bismarck, Plomp) explorent des relations entre l'enveloppe spectrale et les qualités perçues par des auditeurs bénévoles (voir encadré page suivante). Ils mettent en évidence la *brillance* et, dans une moindre mesure la *rugosité* (ou raucité) mais peinent à extraire d'autres paramètres généralisables aux sons synthétiques et stables de ces expériences. L'intérêt se porte alors sur la caractérisation des sons d'instruments de musique.

3.3.3. L'analyse et la synthèse de sons d'instruments de musique

Dans les années 1960, l'acoustique bénéficie des importants et rapides développements des techniques du son (Risset, 1994). Les synthétiseurs analogiques se multiplient et, face à la piètre qualité des registres de timbre censés reproduire le son des instruments réels, des recherches sont entreprises pour caractériser les paramètres acoustiques du timbre causal (Berger, 1964 ; Saldanha et Corso, 1964 ; Clark *et al.*, 1964). La technique de synthèse permet aussi de faire des expériences nouvelles, comme la permutation d'enveloppes spectrales et temporelles (Strong et Clark, 1965). Cependant, les résultats sont complexes, ils dépendent du type d'instrument et ne sont valables que pour les sons d'un synthétiseur donné.

Expériences avec des sons de synthèse : la brillance

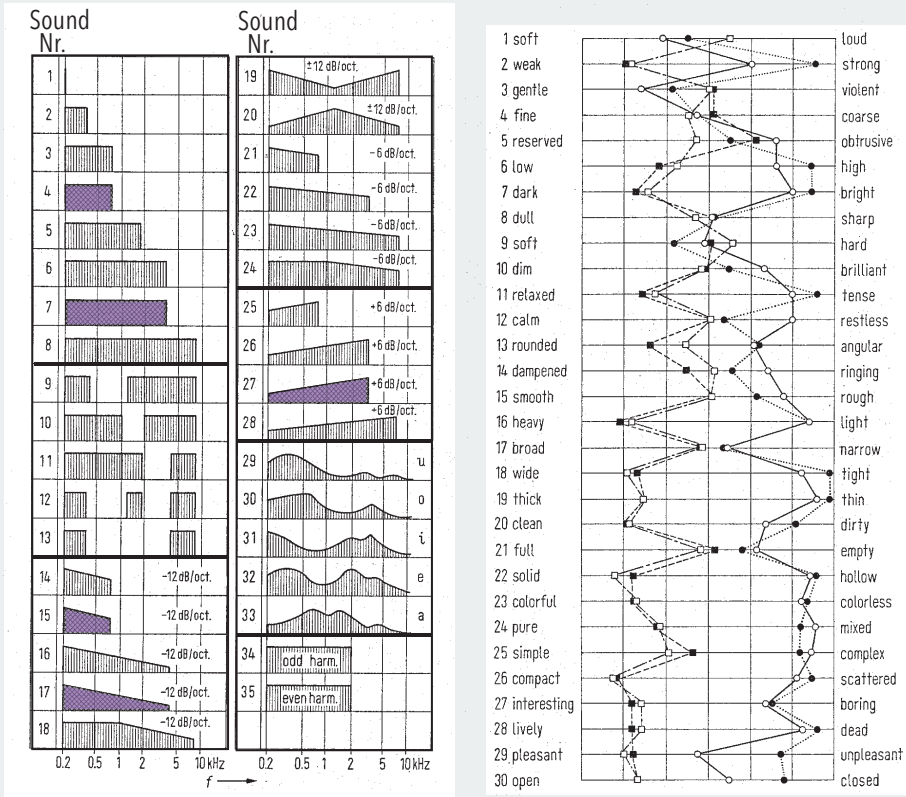


Figure 7.50 Expérience de G. von Bismarck (1974). À gauche : enveloppe spectrale des 35 sons sélectionnés ; les hachures croisées, de couleur violette, indiquent les bruits. À droite : attributs des 30 échelles verbales et résultats pour deux stimuli opposés : cercle pour le son n° 1 (son pur de 200 Hz) ; carré pour le son n° 8 (son complexe de 200 Hz). Les réponses des sujets musiciens sont les points (cercles et carrés) noirs.

D'après von Bismarck, G., 1974, « Timbre of steady sounds », *Acustica*, 30, p. 149.

Lichte (1941) travaille avec des sons de fréquence 130 et 180 Hz, pourvus de 16 harmoniques. Il varie principalement la pente spectrale ainsi que le contenu relatif des harmoniques pairs et impairs. **Bismarck** (1974) travaille avec 35 sons de 200 Hz (figure 7.50, partie gauche) et un grand nombre de paramètres spectraux : nombre d'harmoniques, pente de l'enveloppe, formants vocaliques, « zones de réjection », paramètres qu'il applique également à des bruits continus (en couleur sur la figure).

Les descripteurs verbaux sont utilisés dans les deux études. Lichte demande à ses sujets d'évaluer un écart par rapport à des sons de référence en prenant en compte deux critères : terne et brillant. Bismarck sélectionne 30 adjectifs et met en place des tests d'évaluation selon la technique du différentiel sémantique (figure 7.50 partie droite). Malgré les importantes différences entre ces deux travaux, aussi bien sur les types de sons, leur limite supérieure en fréquence, les techniques de passation des tests, que sur le dépouillement des résultats, il faut remarquer que l'attribut qui permet d'interpréter la plus grande quantité des résultats est commun aux deux recherches : c'est la *brillance*. Pour Lichte,

« brightness » correspond au centre de la distribution de l'énergie le long de l'axe des fréquences[1]. Pour Bismarck, « sharpness » correspond au premier moment du spectre, ou zone de fréquence de la concentration maximale de l'énergie[2]. Les autres attributs sont moins aisément isolables et dépendent des conditions expérimentales[3].

L'étude systématique de l'attribut de brillance montre qu'on peut établir une relation assez simple entre le centre de gravité spectrale et la brillance perçue, sur la base d'un découpage du spectre selon les bandes critiques (voir chapitre 3, § 2.2). Cet attribut dominant aussi bien dans le groupe des sujets musiciens que dans celui des non-musiciens se retrouve dans la plupart des tests d'écoute (Grey, 1977 ; Lakatos, 2000) et se révèle un critère prépondérant dans la classification automatique des sons musicaux.

L'analyse verbale des résultats de Bismarck fait état de nombreuses corrélations entre plusieurs adjectifs en particulier : « high » est fortement corrélé à « bright », « light », « brilliant », « thin » et « clear ». Il en conclut qu'excepté « sharpness » les autres caractères du timbre ne semblent pas pouvoir donner lieu à une description verbale de type psychophysique.

Pour contourner ces difficultés, un autre chercheur, **Plomp** (1967 à 1973), utilise l'analyse multidimensionnelle qui ne nécessite plus l'emploi de mots. Les sujets doivent seulement attribuer une « note » à la plus ou moins grande différence entre deux sons. Un programme mathématique traite ensuite les distances entre les sons et fournit une représentation cartographique visualisant ces distances dans un plan. Plomp utilise des sons synthétiques construits sur la base d'une période extraite d'un son d'instrument de musique et montre que l'espace des réponses des sujets est très similaire à l'espace calculé sur les données acoustiques, c'est-à-dire les spectres des sons du test.

Miller et **Carterette** (1975) effectuent plusieurs tests de dissimilarité avec des sons de synthèse de fréquences différentes, possédant une enveloppe temporelle. Les réponses montrent que les auditeurs discriminent, par ordre d'importance décroissante, la fréquence fondamentale, le caractère percussif ou entretenu de l'enveloppe temporelle, le nombre d'harmoniques (5 versus 3 ou 7).

1. « ... midpoint of energy distribution, along frequency continuum ».
2. « ... frequency location of the orverall energy concentration of the spectrum » (1er moment du spectre).
3. Ce sont : *compactness* qui sépare les sons périodiques et les bruits (Bismarck) ; *roughness* (Lichte) qui caractérise les sons à prédominance impaire.

Une étape majeure est franchie avec le développement de l'ordinateur qui va se révéler l'outil essentiel pour la recherche en perception sonore. Lourde et lente à ses débuts, la synthèse numérique offre plusieurs avantages : les paramètres de la synthèse sont parfaitement contrôlables et le savoir-faire peut se transmettre d'un chercheur à un autre. Enfin, l'analyse et la représentation graphique des sons, de même que le traitement des résultats, s'effectuent avec le même outil. Citons en exemple les études menées à cette époque sur l'analyse et la synthèse des sons de trompette (J.-C. Risset et M. Mathews, 1966) ainsi que sur l'analyse exhaustive des caractéristiques physiques des sons de quatre instruments de la famille des « cuivres » (Luce et Clark, 1967). La numérisation des sons nécessite de gros moyens de calcul, accessibles seulement dans les grands centres de recherche[55], cependant les progrès sont très rapides.

Les tests d'écoute de sons d'instruments de musique « réalistes » et contrôlés se multiplient, d'abord avec des sons numérisés ; ensuite avec des sons synthétisés par modulation de fréquence, puis, dans les années 1980, avec des sons extraits des bases de données publiées. La technique de test la plus commune consiste à former des paires de sons et à demander aux sujets de leur attribuer une note de dissemblance (ou de ressemblance). Elle utilise l'analyse multidimensionnelle[56] ou MDS (MultiDimensionnal Scaling). Cette procédure offre plusieurs avantages. Plutôt que

55. Principalement les Bell Labs (New Jersey) ; le M.I.T., (Cambrige) ; Stanford University, pour les États-Unis. Le CNET, l'IRCAM pour la Fance.

de faire des hypothèses sur les variables physiques pour élaborer un corpus de sons paramétrés, il devient possible de pratiquer des tests avec des sons complexes proches des sons réels, sans faire d'a priori sur les paramètres pertinents – à charge pour le chercheur d'interpréter ensuite les résultats en termes de données perceptives et acoustiques. Le traitement des résultats bénéficie du développement des méthodes numériques d'analyse statistiques des données : analyse en composantes principales (ACP), analyse multidimensionnelle, MDS.

3.3.4. L'expérience de Grey : 16 sons d'instruments différents sur la même note (mib_3)

L'expérience de Grey (1977) est le premier test avec des sons d'instruments de musique. Elle bénéficie d'une double avancée technologique : la numérisation de sons musicaux et le développement de l'analyse MDS. Elle porte sur 16 sons instrumentaux de même hauteur. Afin de garder un contrôle sur les stimuli, les sons, qui proviennent d'enregistrements en studio, sont numérisés[57] de sorte que les partiels puissent être analysés isolément. Chaque son est ensuite reconstitué par synthèse additive avec strictement la même fréquence fondamentale de 311 Hz (mib_3), une durée comprise entre 280 et 400 ms et quelques simplifications des bruits d'attaque et des instabilités de fréquence. Ensuite, l'égalisation en intensité des 16 sons est effectuée à l'oreille. La brève durée et le « lissage » des sons minimisent les variations temporelles inhérentes aux sons réels qui pertuberaient l'écoute. Il n'y a pas de vibrato. Les sujets de l'expérience, 22 musiciens professionnels, doivent attribuer une note de similarité (de 1 à 30) à chaque paire de son. Après un essai portant sur 30 paires, ils évaluent les 240 combinaisons du test.

Avant de donner le détail des sons et de présenter les résultats de cette expérience pionnière, nous proposons d'écouter quelques exemples de ces paires de sons, qui ont été repris ultérieurement dans un grand nombre d'autres expériences : Sons des paires **7.41a**, **7.41b**, **7.41c**, **7.41d**.

Test de John Grey. *(voir page 360)*

Son 7.41a (3″)

Paire 1

Son 7.41b (3″)

Paire 2

Son 7.41c (3″)

Paire 3

Son 7.41d (3″)

Paire 4

Ce test demande beaucoup d'attention, car les signaux sont brefs. On comprend aussi l'importance des essais préliminaires qui permettent de se familiariser avec les sons afin de se constituer une échelle personnelle de « similarité ». Certains auditeurs auront peut-être remarqué aussi qu'ils sont gênés parce qu'ils ne reconnaissent pas toujours les instruments, mais ce n'est pas le propos du test. Il est important d'insister sur ce point.

Les tests avec des sons synthétiques, difficiles à écouter, placent le sujet dans une situation tout à fait artificielle. Par ailleurs, lorsque l'auteur du test propose, comme c'est le cas, d'évaluer la similarité entre des sons issus de sources réelles toutes différentes, les auditeurs sont confrontés à des difficultés supplémentaires dues aux interférences entre leurs connaissances des sources, ici les instruments et les qualités de ces sons. En effet, deux sons de qualités voisines mais provenant d'instruments différents, comme la flûte et le cor, conservent malgré tout une certaine « distance » dans l'esprit de l'auditeur.

Les résultats de ce test appellent plusieurs remarques. Le paramètre qui recueille le plus de poids, celui selon lequel les jugements convergent, est la répartition de l'énergie dans le spectre. Il correspond à la « brillance » déjà révélée lors des premières expériences de Lichte et Bismarck. Il s'agit donc bien d'une dimension psychophysique dont la perception ne semble pas dépendre du type de son. Ensuite, sont invoqués des paramètres temporels que seuls des stimuli évolutifs permettent de révéler. Leur interprétation en est difficile. Compte tenu de ce que

56. Différents algorithmes ont été développés que nous ne pouvons présenter ici.
57. Fréquence d'échantillonnage : 22,5 kHz ; dynamique : 14 bits.

Test de John M. Grey (1977) : note mib_3

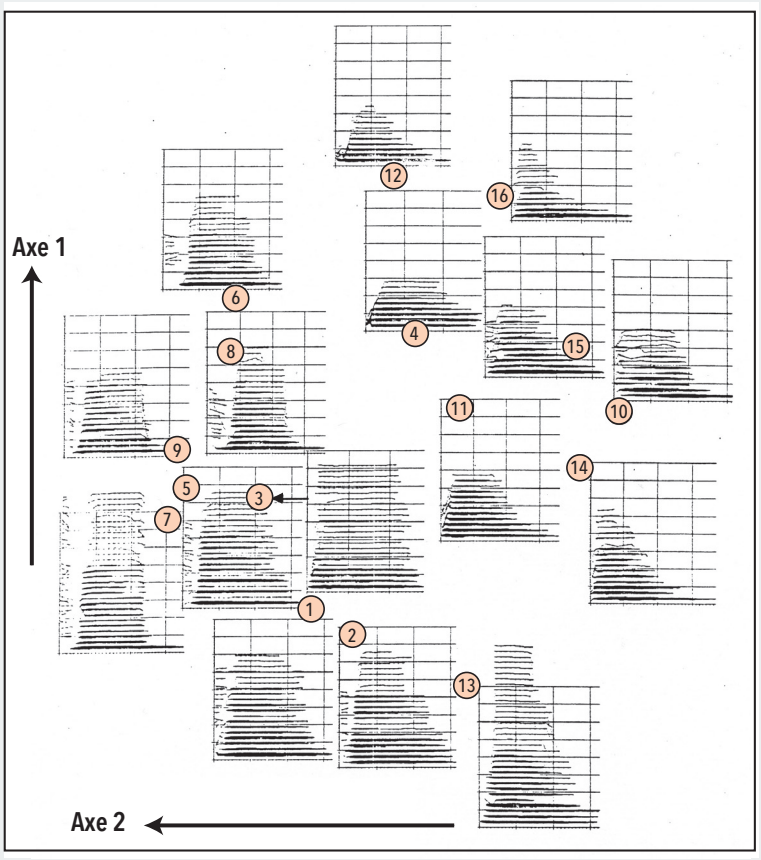


Figure 7.51 Position des stimuli dans le plan des deux premiers axes de l'analyse multidimensionnelle. Les cercles indiquent la position exacte des sons dans le plan de la figure. Les analyses spectrographiques associées permettent d'effectuer une première interprétation acoustique des deux dimensions. Observer en particulier les variations du spectre selon l'axe vertical et les variations du transitoire d'attaque selon l'axe horizontal.

D'après Grey, J., 1975, p. 65.

Les instruments retenus pour le test sont : 1, 2, hautbois (O1 et O2) ; 3, cor anglais (EH) ; 4, basson (BN) ; 5, clarinette en *mib* (C1) ; 6, clarinette basse en *sib* (C2) ; 7, 8, saxophone tenor (X1 *mf* et X2 *p*) ; 9, saxophone soprano X3) ; 10, flûte (FL) ; 11, trompette (TP) ; 12, cor (FH) ; 13, trombone avec sourdine (TM) ; 14, 15, 16 violoncelle avec trois modes de jeu (S1 sul ponticello ; S2 normal ; S3 jeu sur la touche avec sourdine).

Ces stimuli offrent une bonne diversité de sons de qualités différentes : diversité des types d'instruments (11), des modes de jeu, des changements de nuance, et sons de tessitures différentes (clarinette basse et basson jouant dans l'aigu). En toute rigueur, il convient cependant de faire abstraction des noms d'instruments dans la présentation des résultats, c'est pourquoi nous leur avons substitué des numéros. En effet, il ne s'agit pas d'un test

d'identification et nous savons qu'un seul son ne peut représenter le prototype de chaque instrument.

Dépouillement du test. L'analyse des distances produites par les auditeurs conduit à positionner les sons à l'intérieur d'un nuage de *n* dimensions, au sein duquel il s'agit de repérer les axes des dimensions principales, c'est-à-dire ceux qui traversent le plus grand nombre de stimuli. Ces axes (ou dimensions) sont classés par ordre décroissant de « poids ». On procède ensuite à l'interprétation des axes dans les plans qu'ils forment deux à deux.

Dans la dimension verticale (axe 1), les sons pourvus d'un grand nombre d'harmoniques sont situés vers le bas (13, 1, 2) et ceux à spectre réduit qui ont une intensité plus grande dans les fondamentaux (4, 16, 2) vers le haut de la figure.

Dans la dimension horizontale (axe 2), il faut observer la forme du transitoire d'attaque, moins abrupte dans le sens de la flèche, ainsi que l'ordre d'arrivée des composantes[1].

Grey et ses collaborateurs effectuent des analyses quantitatives du contenu spectral et temporel des stimuli et parviennent aux conclusions suivantes :

– l'axe 1 représente la répartition de l'énergie dans le spectre ;

– l'axe 2 concerne la forme du début et de la fin du son, et plus particulièrement le synchronisme des harmoniques supérieurs à l'attaque et à l'extinction (on peut y joindre les fluctuations spectro-temporelles pendant la durée du son) ;

– l'axe 3 (non représenté ici) correspondrait aussi à des critères temporels liés aux transitoires.

1. Rappelons que l'aplatissement des données dans un plan (axes 1 et 2) distord les relations entre les sons qui sont positionnés dans un espace à *n* dimensions.

nous avons remarqué au § 2.3.4, page 311, la perception des transitoires ne semble pas réductible à des dimensions indépendantes ; elle est aussi multidimensionnelle et il n'est donc pas étonnant que les deux axes 2 et 3 s'y rapportent.

Pour éclairer l'interprétation, il est intéressant d'écouter les 16 sons du test, ordonnés selon les valeurs qu'ils prennent sur chacune des dimensions.

Son 7.42 (13″)

Sons classés selon l'axe 1 (voir page 389)

L'écoute du Son **7.42**, qui correspond à l'axe vertical de la figure 7.51, rend bien compte des transformations spectrales progressives qui se produisent depuis le son 13, dont le spectre global, extrêmement riche en harmoniques, contient peu d'énergie dans les fréquences basses, jusqu'au son 12, de spectre réduit et riche en fondamental. La transformation spectrale affecte aussi la qualité du transitoire d'attaque. L'écoute du Son **7.43**, qui correspond à l'axe horizontal, ne fournit pas immédiatement la perception d'une transformation graduelle. Si l'attention se focalise sur le transitoire d'attaque, on remarque que les quatre premiers sons (partie droite de la figure, 10, 14, 15, 16) ont en commun une attaque sans netteté, assez fluctuante et une « note » qui ne s'établit pas franchement. Cette impression change assez abruptement au cinquième son (13), d'attaque très nette, de même que pour les six sons suivants (11, 4, 12, 2, 1, 3) dont le contenu aigu-grave change assez rapidement d'une note à l'autre. Les cinq derniers sons (6, 8, 5, 9, 7) ont pour particularité de commencer sur le fondamental. On peut de la même façon écouter le Son **7.44**, qui correspond à l'ordonnancement des stimuli sur l'axe 3 (non représenté). Des groupements apparaissent, mais on ne perçoit pas de variation graduelle avec la position du son sur l'axe.

Son 7.43 (13″)

Sons classés selon l'axe 2 (voir page 389)

Son 7.44 (13″)

Sons classés selon l'axe 3 (voir page 389)

Les 16 sons du test, de même hauteur et de même intensité, sont différents par leur « timbre » mais aussi par les critères qui les différencient, qui sont de nature variable : instabilité temporelle, qualité vocalique du spectre, type d'attaque (sur le fondamental, l'octave, la quinte), richesse harmonique, netteté de l'attaque, pour n'en citer que quelques-uns.

Autres tests

Les autres tests sur la note *mib*$_3$: la variabilité temporelle

L'expérience pilote de Grey a marqué la recherche en perception du timbre et a suscité une floraison de tests similaires sur la même note (*mib*$_3$ = 311 Hz), avec des variantes dans les sources instrumentales et dans la technique d'élaboration des stimuli. Citons les expériences réalisées avec deux corpus de sons. Le premier comprend 21 sons synthétisés par modulation de fréquence, dont 14 imitent des sons naturels, et 7 sont des hybrides croisant le mode d'excitation d'un type de source (l'archet) avec le contenu spectral d'une autre (spectre du piano)[1]. D'autres auteurs ont utilisé des sons puisés dans la base de données McGill[2] en travaillant sur la même note ou sur des notes voisines. Le fait remarquable est que lorsque des sons impulsifs sont mêlés aux sons entretenus, le premier axe de l'analyse MDS différencie les sons impulsifs des sons entretenus ; le deuxième axe porte la dimension spectrale. Le troisième axe, de faible pertinence, porte selon le corpus de sons et selon les auteurs sur un paramètre temporel mal défini : « flux spectral, *spectral variation*, *time variance* »[3]. Les recherches se sont alors orientées sur la perception des variabilités proprement dites. Les simplifications opérées au moment de la resynthèse avaient déjà permis à Grey & Moorer (1977) de montrer que de nombreuses microfluctuations visibles sur les analyses pouvaient être négligées sans dommage notable pour la perception. Grâce aux progrès des techniques de traitement du signal, plusieurs auteurs (McAdams & Beauchamp, 1999) ont effectué des simplifications systématiques portant sur un grand nombre de paramètres que nous ne pouvons détailler ici. Mais si les discriminations perceptives ont pu être mises en évidence de façon certaine, nous conclurons avec Hajda (2007, p. 262) : « malgré les avancées récentes, la pertinence des paramètres de variation temporelle est loin d'être pleinement comprise ».

Les tests d'écoute : dimensionnel versus catégoriel

La question qui se pose est celle de la validité d'une démarche psychophysique qui fait l'hypothèse de paramètres indépendants pouvant correspondre à des « dimensions » perceptives. Certes, cette technique de test ne fait pas d'hypothèse sur le comportement d'écoute des auditeurs puisqu'ils sont libres de noter une dissemblance entre les sons, mais le fait de grouper les sons deux à deux fait émerger des oppositions particulières, qui changent d'une paire à l'autre, pour le même son. C. Krumhansl a proposé de prendre en compte les spécificités de chaque son (Krumhansl, 1989). Leur interprétation qualitative par Donnadieu (1997, p. 198) correspond bien aux traits identitaires de chaque son, ceux-là mêmes qui disparaissent dans le traitement MDS, du moins tant que le corpus des stimuli ne concerne qu'un nombre limité de sons[4].

Depuis le premier test de 1977, peu d'informations nouvelles ont été produites, malgré la quantité et la qualité des recherches publiées, et il est opportun de s'interroger sur la pertinence d'une méthode d'analyse multidimensionnelle avec des stimuli issus de sources instrumentales différentes. Lorsque le corpus des sons du test croît, (Lakatos, 2000 ; Donnadieu, 1997, figure p. 223), le plan des deux premiers axes fait clairement apparaître le regroupement des stimuli en catégories causales : instruments à vent, cordes libres, percussions de hauteur indéterminée, ce qui signifie que les résultats ne portent plus sur les qualités du timbre mais sur ses aspects identitaires. Nous avons à plusieurs reprises évoqué la primauté de l'écoute événementielle (indiciaire, causale) sur l'écoute des qualités des sons, et surtout sur le fait qu'il s'agit de deux modalités d'écoute différentes, quasi antagonistes, de la perception sonore. Un test qui consiste à sélectionner un son de chaque instrument oriente plus ou moins implicitement sur un traitement cognitif de reconnaissance de la source.

1. Voir les recherches de Krumhansl, 1989 ; Krimphoff, 1994 ; Donnadieu, 1997 ; Faure, 2000.
2. Iverson & Krumhansl, 1993 ; Donnadieu, 1999 ; Lakatos, 2000 ; Hajda, 2007 ; Handel & Erickson, 2001.
3. Voir Hajda, 2007, p. 262.
4. Grey avait déjà remarqué, dans l'interprétation de ses résultats, un regroupement de certains sons par « familles instrumentales ».

De surcroît, plusieurs sons de chaque corpus sont difficilement identifiables, ce qui est explicable. Comment reconnaître à coup sûr un instrument à partir d'un son isolé, surtout dans une tessiture aiguë (basson mib_3), et a fortiori lorsqu'il s'agit de sons modifiés ?[5] Les très intéressants commentaires verbaux recueillis par Anne Faure au cours des tests pratiqués avec le corpus de sons synthétisés révèlent et confirment le fait que les auditeurs s'efforcent avant tout de donner un sens aux sons entendus en décrivant les types d'action pouvant les avoir produits. L'auteur conclut : « les sujets à qui on demande de décrire verbalement les similarités et les différences entre les deux sons entendus décrivent bien souvent non pas le son lui-même mais ce qui aurait pu le produire. » (Faure, 2000, p. 354).

Tous les chercheurs s'accordent pour définir le timbre sur la base de tests d'écoute, en adoptant le même protocole de notation d'une dissemblance entre deux sons. Mais comment interpréter cette note de dissemblance ? Sur quoi porte l'écoute des sujets ? Ces réflexions suggèrent que, pour expérimenter sur l'écoute et la caractérisation des qualités des sons musicaux, il est nécessaire de constituer des corpus de sons relevant de la même catégorie instrumentale. Si, comme nous le suggérons, les qualités musicales sont appréhendées comme des variations, des anamorphoses, des écarts par rapport au prototype partagé par un groupe de musiciens, il nous paraît primordial de ne pas mixer les sources de production et d'offrir à l'auditeur les meilleures conditions de discriminations fines, sensorielles et cognitives.

5. Il est surprenant que les résultats soient toujours présentés dans les articles avec le nom de l'instrument utilisé pour la production du son, comme si ce son particulier constituait un prototype acoustique de l'instrument entier, abstraction faite de toutes les variations dont nous avons vu l'importance pour la constitution même du timbre (tessiture, nuances d'intensité, mode de jeu, etc.).

3.3.5. Un test de timbre qualitatif : un instrument, une note et des transformations paramétrées

Une expérimentation portant sur la comparaison de sons de même hauteur, de même intensité, de même durée et produits par la même source instrumentale permet de focaliser l'écoute des auditeurs sur des aspects très spécifiques de la qualité sonore. La technique de test utilisée est celle de la catégorisation libre, ce qui permet de prendre en compte les stratégies perceptives des sujets. Les deux exemples que nous présentons concernent des sons d'instruments non entretenus : le steel-drum et le piano.

1/ Étude du transitoire d'attaque de sons de steel-drum (P. Gaillard)

Avec cet instrument, la qualité du son dépend des relations de justesse entre les partiels, ainsi que du temps de mise en résonance des différentes parties de la structure. Dans le jeu *mf,* il est fréquent que surgissent, sitôt après la frappe d'une note, des réponses de partiels à l'octave ou à la quinte qui sont dues aux couplages de zones vibrantes connexes de celle qui est frappée (voir Thomas & coll., 2011). Elles peuvent se produire avec un retard important (100 à 200 ms) et une intensité telle que le niveau sonore, qui avait commencé à décroître, remonte et dépasse celui de l'excitation initiale[58].

L'étude réalisée par Pascal Gaillard[59] porte sur un corpus de douze transformations temporelles et spectrales d'un seul son de steel-drum (Son **7.45**). Elle a mis en évidence l'incidence de la durée du transitoire dans l'appréciation de la qualité sonore de cet instrument : certains auditeurs sont capables de suivre temporellement les arrivées individuelles des partiels alors que d'autres les intègrent dans le jugement global de qualité du transitoire ; en conclusion, l'auteur propose de prendre en compte une durée individuelle d'intégration des événements sonores

Son 7.45 (1'18")

Les 13 sons du test steel-drum

58. Les musiciens qui entendent pour la première fois les sons de cet instrument hésitent à dire s'il s'agit d'un son percuté ou entretenu.
59. Pascal Gaillard, 2000 et 2007.

du transitoire. En outre, cette recherche confirme l'intérêt, pour l'expérimentation perceptive, des méthodes de synthèse et de transformation des sons réels numérisés dans toute leur complexité.

2/ Synthèse de sons de piano par paramètres (J. Bensa)

Cinq sons du test piano de Bensa (voir l'encadré ci-contre)

Son 7.46a (6")

2A

Son 7.46b (6")

2D

Son 7.46c (6")

5D

Son 7.46d (6")

5A

Son 7.46e (6")

1B

La modélisation physique des sons de piano a mis en évidence l'incidence de deux paramètres typiques du son de cet instrument : l'*inharmonicité* des partiels et la plus ou moins grande contribution de partiels *fantômes,* résultant du couplage des vibrations transversale et longitudinale[60].

Il est donc possible de modifier par synthèse des sons dont on contrôle indépendamment les contributions de chacun des deux paramètres énoncés, afin de vérifier leur incidence sur la qualité globale du son de piano (voir encadré).

L'exemple du Son **7.46** donne un aperçu des stimuli obtenus par modification des deux paramètres avec le code suivant : de 2 à 5 c'est le paramètre « inharmonicité » qui croît ; de A à D, c'est le paramètre « partiels fantômes » qui croît. Le dernier son, code « 1B », est représentatif d'un son de piano à queue.

Les deux tests que nous venons de présenter portent sur des sons non entretenus dont la composante temporelle est fortement variable. L'interprétation des résultats du test de sons de piano indique que certains auditeurs concentrent leur attention de façon préférentielle sur l'aspect temporel du son (le transitoire d'attaque) et d'autres plutôt sur son aspect spectral (qualités de la note), quelles que soient les modifications effectivement opérées sur les sons.

Il est important de prendre en compte les capacités individuelles d'écoute liées à cette dualité temps/fréquence qui est au cœur de l'analyse acoustique.

Au cours de l'étude du timbre, et de celle plus générale des « qualités » du son, nous avons été confrontés à plusieurs reprises au fait que la perception des sons par les auditeurs reste la grande inconnue des tests.

Il convient maintenant de le prendre prioritairement en compte dans la mise en place du protocole expérimental.

60. Julien Bensa, 2005.

Qualités des sons de piano[1]

Le test comporte 16 sons combinant deux à deux quatre valeurs différentes de deux paramètres : β (inharmonicité) et G (partiels fantômes), plus un son réel témoin. Tous les sons ont la fréquence fondamentale d'un do_1 (62 Hz) et la même enveloppe dynamique globale (attaque et extinction). Le protocole choisi est la *catégorisation libre* (voir chapitre 4, § 4.4) qui permet de faire état de la diversité des orientations perceptives des auditeurs tout en offrant des clés d'interprétation par la verbalisation. Trois groupes de sujets ont effectué le test : des pianistes confirmés, des musiciens non-pianistes et des auditeurs sans compétence musicale déclarée. À l'issue de la formation des catégories, il est demandé à chaque auditeur de caractériser verbalement les différents groupes qu'il a formé. Les 33 sujets qui ont participé à l'expérience ont tous réussi la catégorisation, mais certains ont eu beaucoup de difficultés à caractériser les groupes constitués, car le résultat perceptif est beaucoup plus complexe qu'une simple combinaison de deux paramètres. La seule modification de l'inharmonicité produit au moins deux effets. Quand l'inharmonicité croît, la hauteur perçue varie (le son paraît plus « haut » quand l'inharmonicité augmente) et le son paraît s'enrichir : il devient plus « brillant ». Environ un tiers des auditeurs (12 sur 33) a formé quatre groupes de sons correspondant strictement aux quatre valeurs du coefficient d'inharmonicité.

Les qualifications de ces auditeurs rendent compte de la gradation perçue entre la première et la quatrième valeur de β en utilisant des termes très différents : de terne vers brillant, de grave vers aigu, de bas vers haut, de rond vers métallique, de fermé vers ouvert. Certains d'entre eux font état de sous-catégories, internes aux quatre premières, en mentionnant la présence de partiels aigus dissonants ou des variations de qualité de l'attaque du son lorsque G augmente. Un deuxième tiers des sujets a formé les catégories prioritairement sur les qualités de l'attaque du son : de mou à net, de doux à agressif, mais aussi de sourd à brillant. En effet, l'accroissement du paramètre G produit des effets spectraux similaires à ceux de l'inharmonicité, que certains auditeurs perçoivent de façon plus critique au début du son et attribuent à l'attaque, alors que d'autres les combinent avec l'inharmonicité dans une évaluation globale de qualité spectrale de la note. Enfin, quelques auditeurs du troisième groupe se sont efforcés de prendre en compte les deux critères de qualité d'un son de piano – la percussion et le spectre – mais avec beaucoup de difficultés puisque les deux paramètres β et G ne sont pas indépendants perceptivement.

Les résultats du test font donc état de deux attitudes d'écoute distribuées de façon assez équitable parmi les trois groupes d'auditeurs : l'une donnant la priorité aux aspects spectraux et l'autre à la dynamique temporelle.

1. J. Bensa

4. Le timbre et les pratiques du son

4.1. Les qualités du son : une écoute choisie

« Assis à la terrasse d'un café, j'entends des voix, des bruits de verre, les pas des promeneurs et une quantité d'autres sons que je reconnais en une fraction de seconde et qui, à peine entendus, passent à l'arrière-plan de ma conscience. Et, soudain, je me surprends en train d'écouter la voix d'une personne dont je ne comprends pas la conversation mais dont les intonations me captivent : c'est une voix grave, souple et modulée, harmonieuse, (...) je suis en train d'écouter musicalement cette voix. »

Cette description, que nous avons imaginée, rend compte d'une situation banale que chacun a pu vivre. Soudainement fasciné par un événement sonore, nous nous surprenons à l'écouter intensément, sur une durée de plusieurs secondes, et nous sommes de fait dans l'état d'une d'écoute qualitative. Ce qui est visé est le son pour lui-même et non plus pour ce qu'il signifie (voir chapitre 4). Le changement opéré tient au fait que l'écoute est décidée librement par l'auditeur qui focalise son attention sur des aspects du son qui varient selon ses motivations, ses connaissances, selon sa pratique du son et selon la situation dans laquelle il se trouve. Il existe effectivement une très grande diversité d'écoutes. Un instrumentiste, un facteur d'instrument, un preneur de son, un acousticien, un designer sonore et un compositeur de musique acousmatique réunis dans un même lieu pour évaluer les qualités d'une séquence sonore donnée entendent tous le même signal, mais ils n'écoutent pas la même chose !

Comment rendre compte de cette diversité d'écoutes ? Comment accéder à la description des sons, tels que chacun de nous les entend ? Pour communiquer avec d'autres auditeurs, ou tout simplement pour affiner ses propres perceptions, il faut pouvoir les traduire sous une forme intelligible et si possible objective. Une première démarche consiste à traduire ce que nous percevons des qualités du son en mots, en phrases. Revenons à l'écoute de l'exemple sonore déjà présenté au début du chapitre 4, en engageant vivement le lecteur qui n'aurait pas encore fait l'expérience à s'y prêter avant de poursuivre la lecture.

Son 7.47 (27")

Consigne d'écoute

Vous allez entendre une séquence musicale, veuillez qualifier le son de l'instrument.

L'expérience est à la portée de tout auditeur. Nous l'avons pratiquée pendant plusieurs années, avec des publics variés et en particulier avec trois groupes différents d'étudiants : 1/des musiciens du Conservatoire de Paris (M) ; 2/des étudiants du mastère ATIAM (A)[61] ; 3/des élèves d'écoles d'ingénieurs (I) suivant une option « Sciences-Histoire-Musique » à l'École des mines de Paris, dont beaucoup ont une pratique de la musique. Le dépouillement des textes écrits par les participants a été réalisé à l'aide de méthodes et d'outils linguistiques développés par Danièle Dubois. Seule une analyse complète des discours prenant en compte le nombre et la nature des mots, les tournures grammaticales, la structure des phrases et l'implication du sujet dans l'énonciation permet d'interpréter les processus d'écoute des trois groupes d'auditeurs (Castellengo M. & Dubois D., 2007).

61. ATIAM : Acoustique, Traitement du signal et Informatique appliqués à la musique.

Les premiers résultats de l'analyse confirment tout d'abord l'hypothèse que nous avions formulée : « l'identification précède la qualification ». Malgré la consigne explicite demandant de qualifier le son de l'instrument, tous les participants ont eu comme première préoccupation de chercher à reconnaître l'instrument. Les musiciens (M) s'attardent sur l'identification de l'instrument. Ce sont les plus compétents en la matière et donc les plus désemparés. Dans la phase de qualification, ils s'attachent principalement à décrire les caractéristiques de la musique qui, pour eux, n'est pas dissociée du son. Les auditeurs avertis de la manipulation des sons (A) repèrent rapidement le procédé d'inversion, le signalent et produisent peu de commentaires sur le son proprement dit. Les auditeurs amateurs (I), désorientés par l'étrangeté instrumentale, abandonnent rapidement la tentative de reconnaissance et développent d'abondants commentaires sur le son, sur ce qu'il leur suggère et sur les impressions qu'ils ressentent à son écoute. Ainsi, la même séquence musicale fait sens différemment pour ces auditeurs ; elle est tour à tour le son d'un instrument de musique, un signal sonore enregistré, une séquence sonore étrange qui produit des effets émotionnels divers. Nous concluons ainsi l'étude :

> *En résumé, par la déstabilisation des processus réguliers d'identification et de qualification du timbre musical qu'elle provoque, cette expérience d'un bref extrait passé à l'envers permet d'identifier trois groupes de sujets qui manifestent dans leurs discours trois conceptualisations d'un « même » phénomène acoustique en référence aux constructions cognitives ordinaires qu'ils ont élaborées au cours de leurs expériences et pratiques différentes de la musique et des sons musicaux.*
>
> Castellengo, M., Dubois, D., 2007, p. 34.

C'est bien l'auditeur qui qualifie les sons et c'est lui qu'il faut interroger. Pour envisager une étude expérimentale de l'écoute du timbre qualitatif, il importe donc de prendre en compte les différentes pratiques du sonore, car c'est en leur sein que se développe et s'affine l'écoute, en vue d'un objectif particulier et en référence aux prototypes cognitifs constitués en mémoire.

4.2. La diversité des écoutes dans les pratiques de la qualité sonore

Avant d'être l'objet d'une étude en acoustique, le timbre est un savoir-faire des musiciens et des luthiers. Dans leur pratique instrumentale, ils développent des aptitudes remarquables pour produire des sons tout en contrôlant leurs qualités, ce qui façonne nécessairement l'orientation perceptive de l'écoute, selon « ce qui fait sens » pour chacun d'eux. Aujourd'hui, l'enregistrement du son a pris une place importante dans cette chaîne d'écoute. En fixant le son, l'enregistrement permet une transmission à distance et introduit un décalage temporel parfois important entre le moment de la production et celui de l'écoute. Le développement et la complexification des moyens techniques sont tels que la pratique du son est devenue le fait de spécialistes dont les orientations d'écoute sont aussi très particulières. Un grand nombre de questions se posent, en particulier celle du statut du signal enregistré par rapport aux productions sonores d'origine qu'il est censé reproduire, et celle de la diversité des écoutes. Ces questions sont d'importance pour les acousticiens qui travaillent le plus souvent à partir de sons enregistrés afin de disposer de documents objectifs et reproductibles pour leurs expériences.

4.2.1. Les instrumentistes, les luthiers, les compositeurs

Les instrumentistes et les chanteurs, qui ont le grand privilège d'être au centre d'une boucle « action-perception » fonctionnant en temps réel, sont au premier rang de tous ceux qui créent et transforment les sons. Chacun d'eux développe et renforce des associations entre des gestes précis – instrumentaux ou vocaux – et certaines des caractéristiques sonores sur lesquelles il peut agir. Ce seront, par exemple, l'attaque du son pour un trompettiste, l'interaction complexe entre les qualités de la voix et l'ajustement des voyelles pour un chanteur.

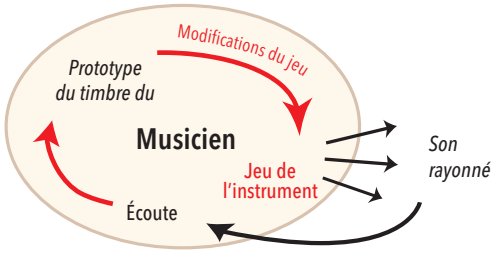


Figure 7.52 La boucle « action-perception-ajustement » d'un instrumentiste.

Luthiers et compositeurs de musique sont dans une situation similaire, à cela près qu'il existe un délai temporel, quelquefois important, entre le moment où ils développent intérieurement leur conception intime des qualités sonores d'un instrument en construction ou d'une partition future et celui de sa « mise en son » effective. Ils doivent donc développer une mémoire exceptionnelle du son adaptée à leur pratique. Pour le compositeur, c'est la mémoire des typologies sonores instrumentales et des effets de leurs combinaisons en fonction de l'écriture musicale. Pour le luthier, c'est celle de l'incidence des transformations apportées à l'objet en cours d'élaboration, comme l'amincissement d'une table ou l'agrandissement d'une perce.

4.2.2. Les pratiques du son enregistré : du timbre au son

Les savoir-faire que nous venons d'évoquer s'inscrivent dans la boucle causale de la production mécanique du son. D'autres acteurs du son travaillent sur un signal sonore enregistré et reproduit par les moyens électroacoustiques : leur écoute se déploie et s'affine en fonction des outils de transformation du son dont ils disposent.

Preneurs de son

Les preneurs de son qui s'inscrivent dans la chaîne traditionnelle de la diffusion et de l'enregistrement d'un concert, par exemple, sont rompus à l'écoute comparée de la production directe et de sa restitution en studio. De même que la photographie n'est qu'une trace réduite de la réalité qu'elle est censée représenter, et ne peut véritablement prendre sens que pour ceux qui possèdent les clés de ce type de représentation, l'enregistrement ne fonctionne pleinement que pour des auditeurs qui ont vécu l'audition directe d'événements similaires. Cette remarque ne concerne pas les interprètes dont l'écoute est généralement orientée, moins sur le « son » que sur le style de jeu, et plus spécifiquement sur leur propre exécution. La prise de son et sa restitution sont des interprétations, des traductions d'une réalité en grande partie insaisissable dans toute sa complexité. L'écoute nécessaire au contrôle des qualités d'une prise de son est multiple et relève autant de l'écoute analytique que

de l'appréciation qualitative globale. C'est une « écoute critique » de l'évaluation du contenu spectral, de l'équilibre des composantes graves et aiguës, de la netteté des transitoires, de l'importance relative des bruits, de l'appréciation de la durée de résonance, etc., autant de critères sur lesquels le preneur de son porte son attention et pour lesquels plusieurs écoutes sont souvent nécessaires. Il doit aussi pouvoir comparer sans cesse le rayonnement réel et sa restitution afin de développer une mémoire auditive spécifique sans laquelle l'appréciation de la qualité sonore n'est pas possible.

Timbre et/ou son

Si les musiciens parlent du timbre de leur instrument, les ingénieurs du son et les compositeurs électroacoustiques utilisent plutôt le terme « son »[62]. Ce terme est également adopté de façon générale par les interprètes qui intègrent l'usage de micros et des différents traitements sonores dans leur prestation : ils « fabriquent » un son de qualité particulière, son qui est aussi leur marque. On notera au passage le transfert sur le terme « son » de l'ambiguïté perceptive identité/qualités. Selon nous, les problèmes perceptifs sont les mêmes, transposés dans une pratique où la chaîne d'écoute est rendue plus complexe par les techniques du son.

Musiciens créateurs de son : informatique ou analogique

Dès les débuts de l'enregistrement sonore, le son matérialisé a fait l'objet de transformations à des fins de création musicale car, dans cette pratique, il n'est plus question de se référer à une production naturelle, à une causalité dont les traces sont le plus souvent honnies (voir § 1.7, page 294). Il faut tout de même des repères pour s'y retrouver, dans un flot sonore indifférencié. Schaeffer et ses collaborateurs se sont attelés à une tâche colossale : l'élaboration d'un nouveau solfège du son écouté pour lui-même, dégagé de toute référence aux sources instrumentales. La systématisation proposée, satisfaisante intellectuellement, ne peut véritablement prendre sens qu'incarnée dans une musique, ce qui s'est révélé difficile. Schaeffer définit une typologie des « objets sonores » sur la base de deux paramètres principaux qu'il nomme la facture et la masse. La facture, qui correspond à la « réalisation instrumentale de l'entretien » se décline selon trois catégories : impulsion, entretien, itération. La masse, qui « correspond de façon large à la hauteur », se présente aussi de trois façons : masse tonique, complexe, variée. En combinant ces deux critères, on obtient donc neuf types d'objets sonores que Schaeffer considère comme « convenables au musical »[63]. Toutefois, le vocabulaire adopté par l'auteur est assez particulier : le terme « impulsion » désigne un son bref, qu'il s'agisse effectivement d'un son percuté ou d'un son entretenu ; le terme « masse » englobe les deux aspects, spectral et tonal de la hauteur. L'étape suivante est celle de la morphologie, qui permet de qualifier la structure des objets sonores ; elle est trop touffue pour être présentée ici[64]. De l'avis même de Pierre Couprie « *le tableau complet est difficilement utilisable car très complexe, ou par moment trop flou* »[65]. Dans la pratique, chaque compositeur forge ses propres outils de description du son[66]. Citons Pierre Henry :

> *Je ne nommerai jamais un son : percuté/lissé, et c'est tout ! Je touve toujours quelque chose de plus littéraire, de plus anecdotique, de plus personnalisé, et donc de plus simple à mémoriser.*
>
> Henry, P., 1979, § 24.

62. Voir Delalande, *Le SON des musiques*, (2001).
63. Schaeffer (1998), *Solfège de l'objet sonore*, réédition en CD des enregistrements de 1967. On écoutera CD III (31-42) les exemples de typologie.
64. Il existe sept critères morphologiques (trois de matière, deux de forme, deux de variation) eux-mêmes déclinés en « types, classes, genres et espèces ». Voir Schaeffer, *Traité des objets musicaux*, p. 584-587).
65. Voir Couprie P., (2001), p. 214.
66. Voir Couprie *op. cit.* ; Denis Smalley et la « spectromorphologie » ; R. Murray Schafer et le « fait sonore » ; François Bayle et « l'image-de-son ».

Ainsi, la boucle « production du son-écoute-appréciation des qualités du son-réajustement des paramètres de la production » est unique pour chaque créateur sonore, non seulement parce que les organes des sens et les capacités d'action diffèrent d'une personne à l'autre, mais surtout parce que l'appréhension individuelle des qualités sonores est extrêmement dépendante de l'expérience d'écoute mémorisée et ne cesse de se transformer tout au long de la pratique. Passé ce constat, il paraît encore plus intéressant d'envisager une voie d'accès aux représentations cognitives du son que chacun de nous développe dans ses pratiques du son et qui sont les référents implicites des évaluations individuelles. Une de ces voies d'accès est l'étude des discours produits par les auditeurs pendant ou après l'écoute.

Parler du son

4.2.3. Transmettre son expérience d'écoute : parler du son

Lorsque le besoin se fait sentir de décrire le son pour communiquer avec d'autres auditeurs, nous tentons de traduire nos sensations en mots. Cette contrainte oblige à appréhender de façon plus claire l'impression ressentie, à la formuler mentalement. Si nous parvenons sans mal à trouver des termes et des expressions pour décrire nos sensations, encore faut-il que nos interlocuteurs en interprètent le sens correctement. Chaque pratique sonore conduit ainsi à développer un jargon, langage spécialisé qui s'affine et se stabilise dans la négociation entre partenaires d'un groupe donné. Les mécaniciens possèdent un vocabulaire assez précis pour décrire le dysfonctionnement d'une machine. La situation est plus compliquée en musique où il faut décrire une impression personnelle, qui est globale et sensible à plusieurs variations concomitantes.

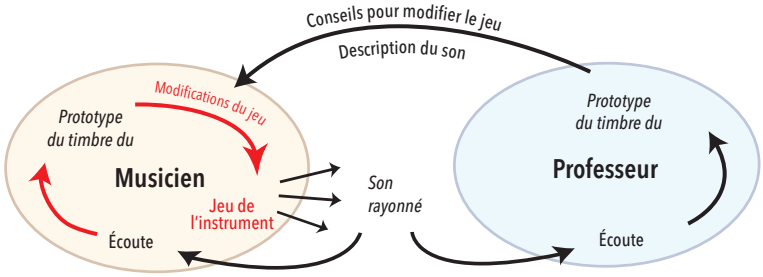


Figure 7.53 Pratique du son et construction d'une « idée du son idéal » dans la relation à l'instrument et dans les interactions avec les autres auditeurs. L'instrumentiste et son professeur n'ont pas nécessairement le même prototype qualitatif du timbre.

Dans la relation pédagogique entre élèves et professeurs, on peut entendre couramment des expressions telles que : « le son doit être plus ouvert, plus clair, il faut lui donner plus de corps, le timbrer, l'enrichir... ». Il peut sembler intéressant de prendre pour point de départ ce langage spécialisé qui correspond à des éléments précis et repérables de la description du timbre tel que le conçoivent les musiciens. Cependant les termes et les expressions employés agrègent plusieurs modalités sensorielles. Ils ne sont pas spécifiques à la conceptualisation sonore et, surtout, ils ne prennent tout leur sens qu'*in situ*, dans l'activité même de la production sonore. Enfin un même terme peut renvoyer à des variations de qualités dont l'interprétation acoustique diffère selon les instruments, comme l'illustre l'étude suivante effectuée sur le son « timbré », expression qu'emploient aussi bien des flûtistes, des clarinettistes que des chanteurs.

4.2.4. Le son « timbré » : un prototype qualitatif dépendant du prototype identitaire

Il est un terme qui revient souvent dans le discours des instrumentistes : le son est « détimbré ». Pour en savoir plus, la question a été posée à deux instrumentistes professionnels – un clarinettiste et un flûtiste – et à trois chanteurs semi-professionnels. Au cours des enregistrements effectués au laboratoire, ces musiciens ont joué (ou chanté) la même phrase mélodique en alternant les deux désignations : son timbré et son détimbré, exactement dans les mêmes conditions de prise de son.

Privilégier une bonne qualité d'écoute pour ces exemples

Écoutons tout d'abord la clarinette (Son **7.48)** en passant du son détimbré au son timbré on remarque tout d'abord un changement global de qualité du son de clarinette, puis, avec un peu d'attention, une petite augmentation d'intensité sur le son timbré, moins de souffle, mais il est plus difficile de décrire le son proprement dit. Chacun peut tenter de le faire à sa façon.

Son 7.48 (12″)

Clarinette : son détimbré puis timbré

La comparaison à l'écoute des sons de flûte détimbré et timbré (Son **7.49**) est plus subtile. Elle se signale également par une diminution du souffle et un changement de qualité sonore assez perceptible mais moins aisé à caractériser par le langage.

Son 7.49 (8″)

Flûte traversière : son détimbré puis timbré

L'analyse spectrographique globale confirme une tendance commune aux deux instruments. Le son qualifié de détimbré (partie supérieure de la figure 7.54) a moins d'harmoniques et plus de souffle, aussi bien pour la clarinette que pour la flûte.

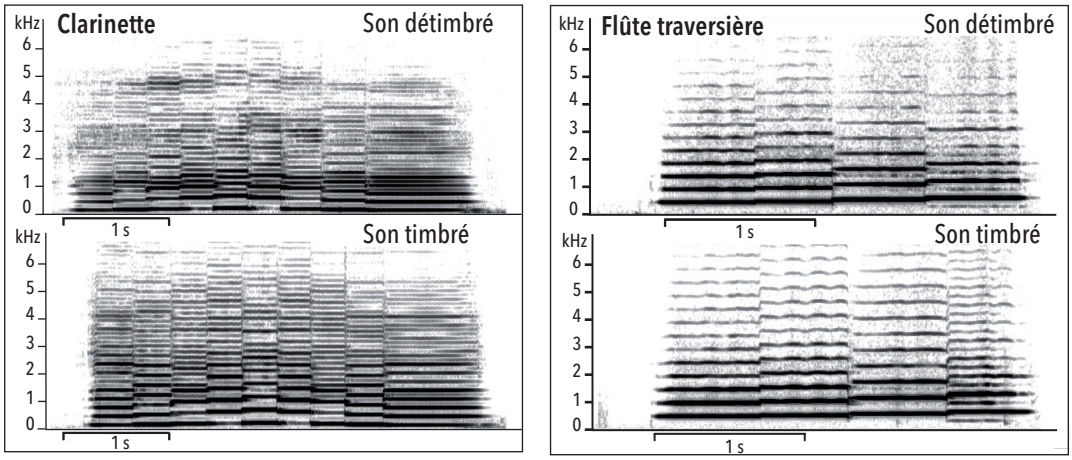


Figure 7.54 Analyses sonagraphiques comparées de jeu avec la qualité « détimbrée » (partie supérieure) opposée à la qualité « timbrée » (partie inférieure). Pour chacun des exemples, la phrase mélodique est jouée ou chantée par le même musicien et enregistrée strictement dans les mêmes conditions.

Pour préciser le changement de sonorité de chaque instrument, nous avons effectué des spectres moyennés sur deux notes : *mib*$_2$ et *lab*$_2$ pour la clarinette (notes réelles) ; *si*$_3$ et *ré*$_4$ pour la flûte traversière (figure 7.55).

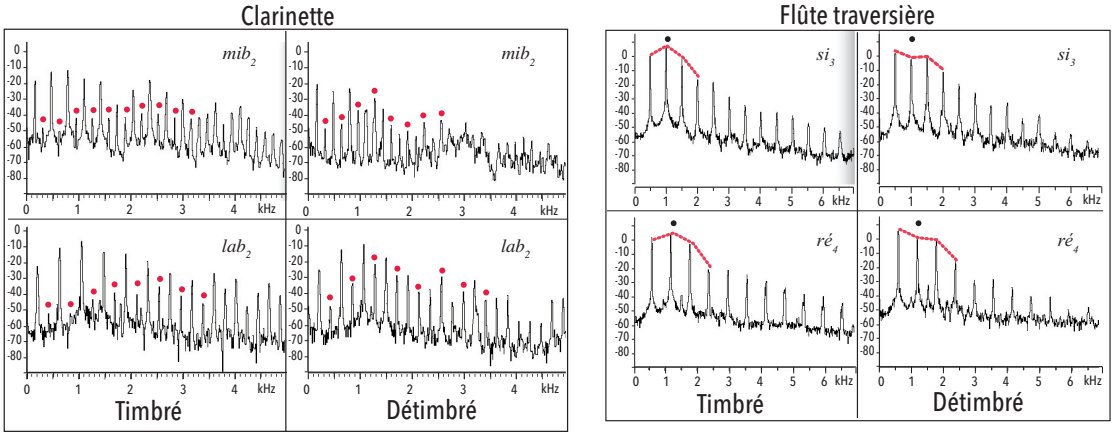


Figure 7.55 Spectres comparés de sons timbrés et détimbrés. Clarinette (1re et 4e note) et flûte traversière (1re et 3e note).

Clarinette. On note, dans le registre grave, la forte prédominance des harmoniques impairs sur les harmoniques pairs (repérés par des points rouges). Il est patent que les harmoniques pairs gagnent en intensité dans les spectres du son détimbré.

Flûte traversière. Il faut observer les changements d'intensité relative entre les deux premiers harmoniques. Dans le son détimbré, le fondamental H1 est le plus intense. Dans le son timbré, l'octave H2 domine. Le tracé rouge qui enveloppe les quatre premiers harmoniques rend sensible ce changement sur les deux notes, *si*$_3$ et *ré*$_4$. Les variations sont notables, de l'ordre de 5 à 10 dB.

Pour ces deux instruments, on observe des modifications communes aux deux sources – changements d'intensité, du nombre d'harmoniques et du souffle – et d'autres qui sont spécifiques à chaque instrument. Qu'en est-il de la voix ?

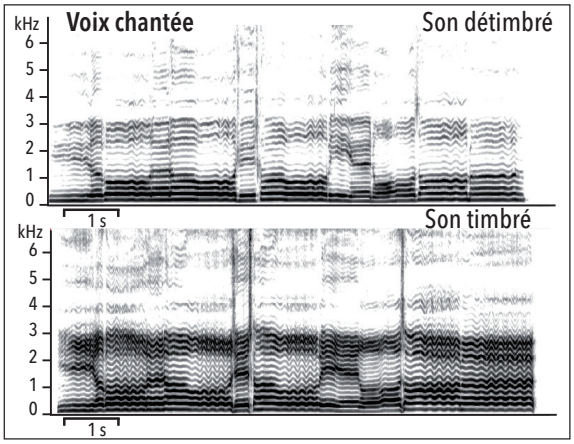


Figure 7.56 Analyses sonagraphiques comparées du même chanteur (baryton) produisant deux qualités de voix : détimbrée et timbrée. Notez les changements du vibrato et de l'intensité du formant du chanteur.

La comparaison des deux phrases chantées avec une qualité détimbrée puis timbrée (Son **7.50**), montre des résultats similaires : augmentation du souffle et du nombre d'harmoniques. À l'écoute, on remarque que les variations d'intensité y sont beaucoup plus importantes que pour les instruments. Concernant les caractéristiques propres, le chanteur *détimbre* en réduisant fortement le vibrato et en atténuant l'intensité du formant situé vers 3 kHz, dit « formant du chanteur », qui est la marque spectrale d'une voix travaillée pour l'opéra. Ces changements sont communs à la plupart des chanteurs (Garnier, 2003).

Son 7.50 (22″)

Voix chantée : son détimbré puis timbré

Le tableau ci-dessous rassemble les observations pour les trois sources.

Tableau 1. Quand on *détimbre* le son

Instrument ou voix	Nombre d'harmoniques	Intensité moyenne	Souffle	Vibrato	Changement spectral	Autre
Flûte	diminue	- 0,5 dB	augmente	sans changement	H1 domine	---
Clarinette	diminue	- 2 dB	augmente	---	H pairs plus intenses	---
Voix chantée	diminue	- 5 dB	variable	diminue ou s'arrête	formant du chanteur plus faible	articulation de type « parole »

En conclusion, les variations de qualité de ces trois exemples d'instruments à vent se traduisent :

- pour partie par des changements communs aux trois sources : le rapport son périodique sur bruit, le nombre de composantes, les variations d'intensité ;
- pour partie par des changements caractéristiques de chaque instrument.

En *timbrant,* le musicien accentue la typicité sonore de l'instrument : il joue sur les « champs de liberté » des paramètres de réglage du son dont il disposte, paramètres qui diffèrent d'un instrument à l'autre. L'attribut timbré, qui correspond à l'intention d'un instrumentiste de donner au son plus de consistance, plus de richesse, plus de « corps », se traduit par des indices acoustiques différents selon les sources sonores : chaque interprète accentue les caractéristiques du prototype identitaire. Catégorisation identitaire et catégorisation qualitative sont ici étroitement liées.

4.2.5. La description vocalique des qualités spectrales

La hauteur spectrale est une dimension perceptive de toute production sonore. Plusieurs chercheurs se sont attachés à l'exploration de la dimension spectrale pour mettre en relation des zones du spectre et des caractéristiques perceptives. Nous avons déjà vu au § 3.3.2 l'importance de l'attribut de brillance, lié à au renforcement de l'énergie dans la zone spectrale du maximum de sensibilité de l'oreille (3 kHz), dont la pertinence est indépendante de la catégorie instrumentale et de la fréquence fondamentale pour les sons périodiques.

Nous présentons, dans l'encadré de la page suivante, les travaux de chercheurs qui explorent la dimension spectrale et proposent une caractérisation phonologique des qualités du timbre. Pour rendre compte du timbre qualitatif dont la perception et la description diffèrent selon les groupes d'auditeurs, il importe d'introduire ceux-ci dans la boucle expérimentale. Nous proposons, dans ce qui suit, la présentation d'un environnement méthodologique dit « sémioacoustique » (pour plus de détails voir Dubois, 2009, chapitre 1).

Caractérisation « vocalique » des qualités sonores

Les voyelles sont des timbres que nous reconnaissons aisément à la position spectrale de leurs formants. La référence aux voyelles peut guider utilement l'écoute de sons possédant des formants, pour peu que l'on garde à l'esprit qu'il s'agit d'une écoute catégorielle dont la structure est liée au système phonologique d'une langue donnée. Il faut aussi apprendre à repérer quel formant guide l'écoute : F1 ou F2. Ainsi, en écoutant l'exemple sonore **7.51** que nous avons reproduit d'après Schouten[1], on entend tout d'abord des voyelles, mais peut-être aussi une mélodie[2], celle du deuxième formant que l'on peut suivre sur la figure. Voyelles, mélodie, ce n'est pas encore une écoute spectrale pure. Cela est-il possible ?

Son 7.51 (21")

Séquence de voyelles chuchotées, qui peut évoquer une mélodie

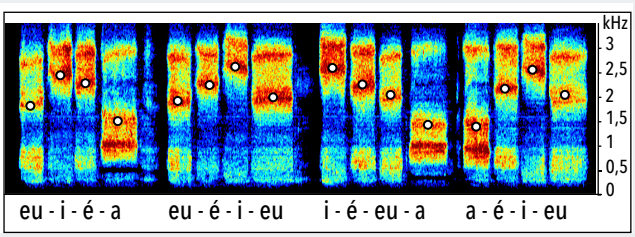


Figure 7.57 Sonagramme d'une séquence chuchotée. Voyelles ou mélodie spectrale ?

Deux auteurs-compositeurs proposent de transposer l'expertise d'écoute spectrale des voyelles à l'ensemble des sons, afin de définir des qualités sonores basées sur la seule répartition spectrale de l'énergie. Ils s'inspirent de la théorie des oppositions phonologiques développée par Roman Jakobson[3].

Slawson (1985) caractérise la « couleur sonore » sur la base de quatre paramètres indépendants de la fréquence fondamentale : « acuteness ; openness ; smallness ; laxness ». Il énonce des règles de perception des variations spectrales valables pour toute source sonore et réalise des exemples sonores illustrant sa théorie. Certains de ces exemples, parfois très rustiques, posent de difficiles problèmes d'écoute en particulier lorsque les formants sont voisins de ceux d'une voyelle, car la référence au système linguistique entre en compétition avec les repères habituels d'une écoute musicale, surtout lorsque l'exemple est construit avec des sons possédant une hauteur tonale.

Cogan (1984) part de façon similaire des données vocaliques pour proposer une grille d'analyse du son des musiques. Il conçoit et réalise un analyseur analogue au sonagraphe, au moyen duquel il obtient la représentation visuelle du son enregistré selon une échelle logarithmique de huit octaves. Il procède ensuite à une description formelle du contenu acoustique sur la base d'une échelle de 13 oppositions par lesquelles il met en évidence les relations énergétiques entre différentes zones de la tessiture et leur évolution en fonction du temps. Les analyses portent sur des comparaisons d'instruments (pianoforte comparé au piano moderne) sur des orchestrations, ou sur des musiques de synthèse dont l'organisation spectrale se prête particulièrement bien à ce type d'investigation.

1. Schouten, 1962, p. 195-207. Voir aussi Erickson, 1975, page 34, qui détaille les voyelles qu'il a choisies pour produire cette « mélodie » chuchotée.
2. C'est « le carillon de Westminster ». qu'on peut noter en partant arbitrairement d'un *do : do mi ré sol, do ré mi do, mi ré do sol, sol ré mi do.*
3. Jakobson décrit les systèmes phonologiques de toutes les langues du monde sur la base de douze oppositions : neuf oppositions de sonorité (vocalique/non vocalique, consonantique/non consonantique, compact/diffus, tendu/lâche, voisé/non voisé, nasal/oral, discontinu/continu, strident/mat, bloqué/non bloqué) et trois oppositions tonales (grave/aigu, bémolisé/non bémolisé, diésé/non diésé), Roman Jakobson, 1963, p. 128.

C. Traube et N. d'Alessandro s'inspirent des recherches de Slawson pour caractériser les changements de sonorité de la guitare selon la position du point de pincement et les différents modes d'attaque. Ils établissent un parallèle entre les onomatopées, qui permettent d'imiter les sons de façon analogue à celles qui sont utilisées dans l'apprentissage du tabla, et les variations de sonorité de la guitare qu'ils formalisent selon les concepts phonologiques proposés par Slawson : relâché/tendu, compact/diffus.

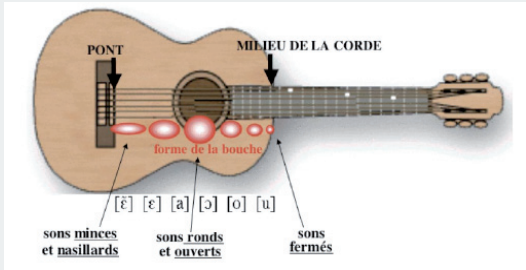


Figure 7.58 Gestes phonétiques sous-jacents à la description du timbre de la guitare classique et voyelles correspondant à l'imitation vocale de sons de guitare joués à différentes distances du chevalet.
D'après Traube, C., & D'Alessandro, N., 2005.

4.3. Une méthode d'étude en sémioacoustique

Plusieurs chercheurs ont proposé d'aborder l'étude du timbre en prenant pour point de départ les conceptualisations des auditeurs, « ce qui fait sens » pour eux. L'approche sémioacoustique, développée au LAM depuis 1995, privilégie les stratégies *top-down* (voir chapitre 4, figure 4.46).

La définition et la mise en œuvre des étapes d'une expérimentation diffère quelque peu selon que l'étude concerne l'évaluation de sons enregistrés ou celle d'instruments essayés en situation de jeu mais, dans tous les cas, le point de départ est l'avis des auditeurs (instrumentistes, luthiers) et les sons proviennent d'une source sonore identifiée (instrument de musique, moteur, etc.).

L'organigramme de la figure 7.59 se rapporte à un cas d'école : l'évaluation comparée de 23 aspirateurs du commerce pour lesquels seul le niveau global d'intensité sonore en décibels est pris en compte habituellement. L'étude a permis d'identifier plusieurs critères spectraux qui rendent compte des différences dans les jugements d'agrément observés au sein d'un groupe de machines égalisées en sonie (Guyot, Castellengo, 1996). Les étapes de cette étude ont été adaptées aux instruments de musique.

- Étape 1. Constitution et sélection du corpus de sons, enregistrements et mise en place du test d'écoute

Il est important de placer les acteurs du test dans les meilleures conditions possibles, de reproduire au mieux une situation d'écoute ou d'interaction réaliste (manipulation, jeu de l'instrument). L'enregistrement d'extraits sonores devant servir aux tests d'écoute implique de faire plusieurs choix : celui du type de prise de son (dispositif de microphones, position et distance par rapport à l'instrument) et celui du lieu d'enregistrement puisque le musicien ajuste son jeu en fonction du retour qu'il reçoit. Il faut ensuite décider du contenu sonore à capter, qui doit être à la fois représentatif de toutes les caractéristiques sonores de l'instrument et concentré dans une durée brève pour se prêter à l'écoute comparée de plusieurs sources dans un temps raisonnable, afin d'éviter la fatigue des auditeurs. Dans le cas idéal où l'on peut disposer de plusieurs exemplaires d'un même instrument, le luthier ou l'instrumentiste doit pouvoir procéder aux manipulations et aux formules de jeu qu'il pratique habituellement : elles sont le condensé d'une longue expérience pour repérer les données les plus sensibles, tant pour le fonctionnement mécanique que pour l'écoute.

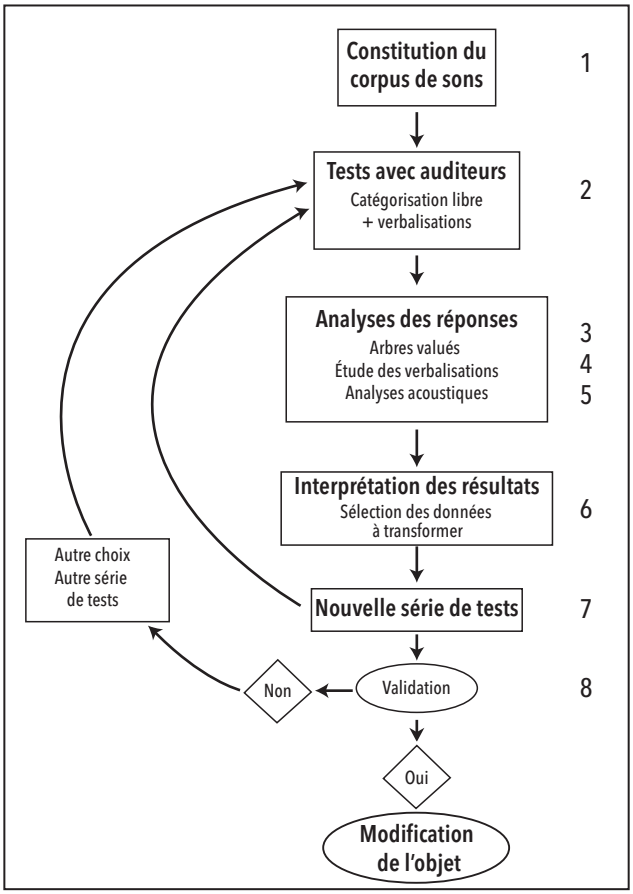


Figure 7.59 Organigramme d'une étude de qualité perceptive.

- Étape 2. Test de catégorisation libre et verbalisations

C'est la pierre angulaire de la méthode. Au cours d'un test avec des sons enregistrés, les auditeurs ont accès librement à l'ensemble des sons qu'ils doivent évaluer. Il leur est demandé de former des groupes selon leur convenance, sans limitation de nombre ni de durée. À l'issue de cette tâche, chaque participant est invité à exprimer verbalement les raisons qui l'ont conduit à effectuer les choix et les regroupements constatés. Pendant les verbalisations, il peut réécouter les sons ou répéter un essai pour préciser sa pensée mais sans modifier en aucune façon le classement auquel il a abouti. Il existe plusieurs logiciels qui permettent d'organiser le déroulement du test et le recueil des résultats. Celui développé par P. Gaillard[67] est conçu pour se coupler de façon optimale aux dépouillements de l'analyse des distances entre sujets et entre catégories ainsi qu'au recueil des verbalisations.

- Étape 3. Analyse statistique des groupements et construction d'arbres valués

Nous ne pouvons pas, dans le cadre de cet ouvrage, présenter plus en détail l'analyse des résultats par la méthode des arbres valués dont un exemple a déjà été donné au chapitre 4, § 4.3.1. Le lecteur peut se reporter aux ouvrages de D. Dubois, Barthélémy et Guénoche cités en bibliographie et, pour une première approche à Gaillard (2009).

67. Voir Gaillard (2009) pour une documentation et http://petra.univ-tlse2.fr/tcl-labx/spip.php ?rubrique2 pour le logiciel.

- Étape 4. Analyse linguistique des données verbales

Le recueil et plus essentiellement l'analyse des verbalisations, a priori individuelles et subjectives, sont de la compétence des linguistes qui ont développé des méthodes appropriées et exigeantes pour traiter cette matière.

> *Il s'agit donc, – si on s'inspire de la métaphore des sciences de la nature qui, depuis bien longtemps, se sont dotées d'instruments de mesure permettant d'arpenter le monde –, de mettre au point et de régler les analyses linguistiques afin de leur faire acquérir la « sensibilité » d'instruments de mesure du subjectif qui soient valides, fiables, en regard des spécificités des objets mesurés.*
>
> Dubois, D., 2009, p. 32.

L'analyse ne s'en tient pas seulement à l'établissement de listes de mots mais à leur interprétation, en fonction du contexte des phrases et des expressions, aussi frustes soient-elles. Citons l'exemple particulièrement frappant, figure 7.60, de l'étude des différents sens que prend le terme « clair » dans l'expression parlée de pianistes invités à évaluer les qualités de neuf pianos de concert (Gherghinoiou, 2005).

Ainsi conçue, l'analyse linguistique permet de remonter aux conceptualisations des auditeurs pour permettre, *in fine*, d'établir des correspondances avec l'analyse acoustique[68].

- Étape 5. Étude des paramètres acoustiques corrélés aux catégories et aux verbalisations

Les données du test de catégorisation et celles des verbalisations permettent de sélectionner, parmi les analyses acoustiques, celles qui offrent les résultats les plus pertinents au regard de la perception.

- Étape 6. Sur la base de l'interprétation des trois types de données – catégorisation, verbalisations, analyses acoustiques, se décide la sélection des paramètres acoustiques à modifier par traitement de signal afin de valider les résultats.

- Étape 7 – Le bouclage

La transformation des sons est suivie d'une nouvelle série de tests incluant les sons transformés sur la base des résultats des analyses antérieures.

Les étapes 6 et 7 sont essentielles pour valider la pertinence de l'étude au regard de la perception. Ainsi, lors de l'étude de la qualité acoustique des sons d'aspirateurs, nous avions formulé l'hypothèse que la présence de fréquences stables produisant une note pouvait être gênante. Quelques sons ont été transformés par analyse-synthèse et, lors d'un nouveau test de catégorisation, les sons transformés ont été regroupés avec les sons acceptables (Guyot, 1996).

- Étape 8. Analyse des résultats et validation de la pertinence des paramètres identifiés

À l'issue d'une nouvelle série de tests et de l'analyse des résultats, une validation négative conduit les chercheurs à s'interroger sur la pertinence des paramètres sélectionnés pour la mise en place du test au regard de la question initiale. Une validation positive donne les orientations recherchées pour transformer l'instrument ou la machine (Le Nindre, 2006).

68. Il est particulièrement éclairant de consulter à ce propos les verbalisations, même sommaires, dont font mention Donnadieu (1997) et Faure (2000). Effectuées a posteriori, après les tests MDS, leur analyse met sérieusement en question l'interprétation « objective » des axes.

Vous avez dit « clair » ?

L'analyse des verbalisations fournies par des pianistes comparant les qualités de neuf pianos de concert a permis de montrer la polysémie particulièrement intéressante du terme « clair » dont l'occurrence est extrêmement courante dans le vocabulaire de ces instrumentistes. À l'issue d'une étude complète du contexte syntaxique (voir tableau joint), Pascale Cheminée montre que le mot « clair » peut endosser au moins trois sens différents dans le corpus étudié.

1/Clair = défini, par opposition à pâteux, flou, brouillon.

2/Clair = sec, dur, par opposition à chaud, rond.

3/Clair = lumineux, transparent, par opposition à sombre.

Connotation	**Positive**	**Positive**	**Négative**
Adjectif	**Clair = défini**	**Clair = lumineux**	**Clair = sec**
Adjectifs associés	Net	Brillant	Brillant
	Précis	Riche	Percussif
	Bien détaché	Profond	Percutant
	Défini	Incisif	Métallique
		Cuivré	Sec
		Chaud	Vert
		Rond	Clinquant
		Ouvert	Perçant
			Dur
			Aigu
			Criard
			Agressif
Contexte syntagmatique	Qui a une définition très bonne, à la fois très *clair*. Quand je dis *clair*, on a une audition facile des notes.	Le plus *clair*, il a vraiment une certaine transparence.	Qui n'a pas de chair dans le son, c'est perçant et *clair* ; ça perce l'oreille ; peu riche harmoniquement.
Antonymes	*Imprécis, brouillon, sourd, pâteux, confus.*		*Profond, riche ; chaud, rond, doux*
Emploi adverbial		Ça sonne *clair*.	

Figure 7.60 Tableau des trois sémantismes de l'adjectif *clair* d'après les contextes dans lesquels il apparaît et les adjectifs auxquels il est associé.

D'après Cheminée, P., 2009, p. 332. L'expérience acoustique sur les pianos a été réalisée par Charles Besnainou (Gherghinoiou, 2005) et Danièle Dubois.

4.4. Les problèmes posés par l'évaluation qualitative des instruments de musique et de la voix

4.4.1. Le prototype de référence de qualité d'un instrument de musique

La capacité à évaluer les qualités sonores d'un instrument de musique est une expertise qui s'acquiert par la pratique d'instruments différents dans une même catégorie. Un interprète évalue les possibilités de jeu : facilité, variations de sonorité, souplesse, malléabilité spectrale. Il s'assure de pouvoir produire le son qu'il désire, sans trop d'effort et avec le contrôle le plus sûr. En effet, l'évaluation des qualités qui s'inscrit dans l'interaction instrument-instrumentiste, est loin d'être exclusivement sonore. Une partie implique l'audition ; une autre concerne le toucher, la vue, le plaisir global du contact avec l'instrument. Finalement, le prototype de qualité qui sert de référence à l'évaluation relève essentiellement du vécu personnel et intègre les estimations issues des comparaisons entre différents instruments (voir chapitre 4, figure 4.44).

En interrogeant un hautboïste professionnel, nous avons obtenu les réponses suivantes pour les qualités d'un bon hautbois : 1/ la sonorité : fluidité, élasticité dans le son, rondeur du son, homogénéité, ce qui signifie globalement un instrument souple qui suit le musicien ; 2/ la justesse ; 3/ le volume ; 4/ la facilité de jeu. Il est à parier qu'un autre instrumentiste aurait des avis différents[69].

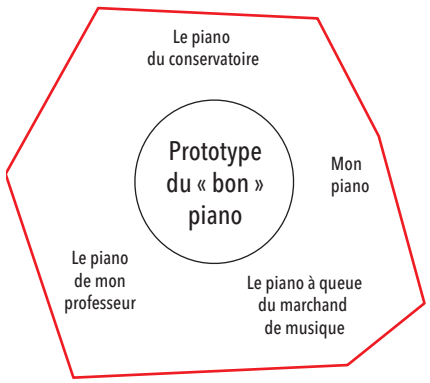


Figure 7.61 Mémoriser toutes les possibilités sonores des sons du piano pour se constituer des références propres à l'instrument considéré, et pouvoir évaluer les qualités d'un instrument de même catégorie.

4.4.2. Les précautions à prendre lors de la mise en place de tests d'évaluation

Pour mettre en place un test d'écoute afin de comparer la qualité de différents instruments, il est nécessaire de respecter plusieurs conditions qui peuvent s'avérer contraignantes.

- Le diapason et l'accord. Tous les instruments à comparer doivent être strictement accordés au même diapason et, pour les instruments polyphoniques, au même tempérament, car la sensation globale de sonorité est dépendante du diapason et du système d'accord.
- L'instrumentiste. Il est nécessaire qu'un seul instrumentiste joue tous les instruments à comparer, après s'être familiarisé avec chacun d'eux. L'adaptation réciproque qui se noue fait partie du jeu.
- Le lieu d'écoute et d'enregistrement. C'est un point important, car nos références de qualité sont celles de sons entendus dans des salles de musique ou des salles de concert, mais certainement pas en chambre anéchoïque. De plus,

69. Lire à ce propos le compte-rendu d'un échange entre instrumentistes et facteurs de différents instruments (Collectif-ITEMM, 2004) ainsi que l'interview du facteur de clavecins Marc Ducornet (Dumoulin, N., 1996).

les musiciens qui « font le son » (violoniste, flûtiste et chanteur) adaptent leur émission selon le retour sonore de la salle.

- La prise de son. Tout en s'attachant les conseils d'un preneur de son professionnel qui aidera à trouver une position pour rendre compte au mieux du rayonnement de l'instrument, il importe d'être suffisamment proche pour capter les bruits caractéristiques de l'instrument afin d'obtenir de bonnes analyses des transitoires[70]. La recherche du bon placement des microphones dépend de chaque lieu et de chaque instrument.
- Le programme sonore. C'est le point le plus difficile à décider. Du point de vue organologique, le seul fait de jouer tous les sons les uns après les autres, en gamme chromatique, fournit déjà une grande quantité d'informations assez faciles à exploiter en analyse acoustique. Ce programme convient aux instruments à clavier, constitués de modules indépendants pour chaque « note » (orgue, clavecin, piano). Cependant, l'écoute est inintéressante, excepté pour les facteurs qui pratiquent ce test en cours de construction. Les instrumentistes et les musiciens auditeurs ont développé leurs références de qualité à l'écoute de musiques qui exploitent les ressources expressives de l'instrument. Il faut donc sélectionner un programme d'extraits musicaux courts, entre 15 et 30 secondes, donnant à entendre différentes tessitures et modes de jeu caractéristiques. Les gammes et la musique fournissent des données complémentaires.

La démarche adoptée pour ce type d'étude implique une bonne coopération entre plusieurs acteurs issus de cultures différentes, ce qui nécessite de prendre le temps d'établir un dialogue intelligible et confiant entre tous les participants et de donner une large part à l'échange verbal, car la perception des qualités d'un instrument ne se restreint pas au son. Entrent en ligne de compte la vue et le toucher, le confort de jeu, et ce rapport complexe entre un instrumentiste et l'instrument dont il rêve, celui qui lui offrirait toutes les possibilités de s'exprimer librement, tout en conservant certaines résistances qui sont pour lui le sel de sa pratique (voir Collectif-ITEMM, 2007).

Plusieurs études conduites au LAM ont bénéficié des compétences croisées entre acousticiens, linguistes, facteurs, instrumentistes, et ont donné lieu à des rapports de recherche consultables en bibliothèque. Les cas de figure sont extrêmement divers :

1. Comparaison du « timbre » de neuf guitares (G. de Montchalin, 1995)
2. Étude de l'identité d'un orgue avant et après restauration (Ch. Piron, 1993)
3. Étude acoustique de la qualité sonore de jeux d'anches de l'orgue Saint-Bénigne de Dijon, par comparaison avec les jeux correspondants des orgues de Dole, de Poitiers et de Saint-Maximin (M. Castellengo, P. Goad, 1996)
4. Étude de la qualité sonore d'un ensemble de clavecins (N. Dumoulin, 1996)
5. Perception de la qualité des archets de violon (B. Saint-Loubry, 1997)
6. Évaluation de la qualité de neuf pianos (Ch. Besnainou, C. Gherghinoiou, 2005)

4.4.3. Les qualités d'une voix chantée

Qui écoute une voix écoute une personne. Le cas particulier de la voix chantée est sans doute le plus passionnant mais aussi le plus complexe à étudier. Ni le chanteur, ni l'auditeur n'ont accès à l'observation directe du fonctionnement de la source. Tous les indices perçus (sur la base desquels s'élabore une interprétation) sont le résultat d'un travail impliquant le ressenti proprioceptif (sensations physiologiques personnelles plus ou moins confuses) et la perception sonore qui, pour le chanteur, combine l'effet des vibrations affectant directement l'oreille interne par transmission osseuse et les vibrations aériennes du son proprement dit (tel que le reçoit le professeur). De son côté, le professeur interprète ce qu'il entend sur la base

70. Les résonances de la salle, que nous sommes capables d'abstraire à l'écoute, perturbent irrémédiablement les analyses spectrographiques.

de son expérience personnelle de la pratique vocale et par référence à l'image mentale personnelle de ce qu'il définit comme une « bonne technique ».

Les discours sur la voix et sur les techniques vocales sont très riches mais difficilement interprétables tant ils sont marqués par l'expérience individuelle de chacun.

Le diagramme de la figure 7.62 élaboré par Maeva Garnier présente une synthèse évocatrice de la complexité des domaines qui interfèrent, à des degrés divers, et qu'il faut prendre en compte lors d'une recherche acoustique sur le timbre vocal. Les caractéristiques sonores, abondamment détaillés dans la figure, ne peuvent être séparées de la technique, de la physiologie. Elles doivent être interprétées différemment selon les types de voix et en fonction des effets spécifiques à un style de chant.

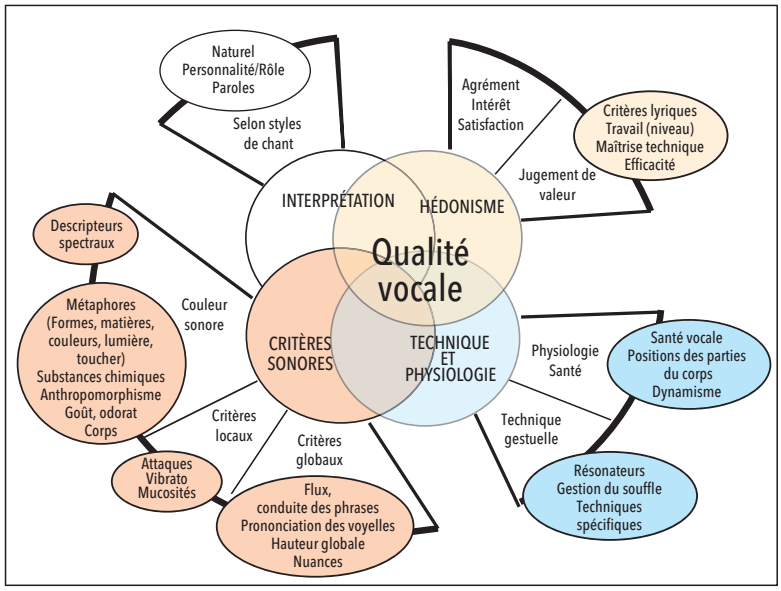


Figure 7.62 Interactions entre les différents domaines qui concourent à la formation des qualités d'une voix chantée.

D'après Garnier *et al.*, 2005.

Plus que pour un instrument de musique, l'implication des chanteurs est primordiale à toutes les étapes de la recherche, de même que celle des linguistes.

Parmi les études développées au LAM citons :

- *Approche de la qualité vocale dans le chant lyrique : perception, verbalisation et corrélats acoustiques* (Maeva Garnier, 2003),
- *Perception et verbalisation de la qualité vocale dans le chant lyrique occidental* (Henrich, N., Bezard, P., Expert, R., Garnier, M., Guérin, C., Pillot, C., *et al.*, 2007),
- *Caractérisation de la qualité vocale dans le chant lyrique occidental : des jugements des professeurs de chant aux descriptions acoustiques.* (Garnier, M., Henrich, N., Castellengo, M., Sotiropoulos, D., & Dubois, D., 2007).

4.5. Les étapes de la construction acoustique du timbre

Le figure 7.63 propose un tableau très schématique des différentes étapes de la chaîne acoustique du timbre d'un instrument de musique. Le chercheur acousticien, situé au dernier maillon, n'en recueille qu'une instanciation bien particulière sous forme d'un signal sonore à analyser.

L'instrument de musique est un système mécanique dans lequel l'interaction excitateur/résonateur conditionne la forme spectrotemporelle des sons émis, donc l'identité acoustique prototypique des catégories de base : la flûte, le piano, le violon, la guitare.

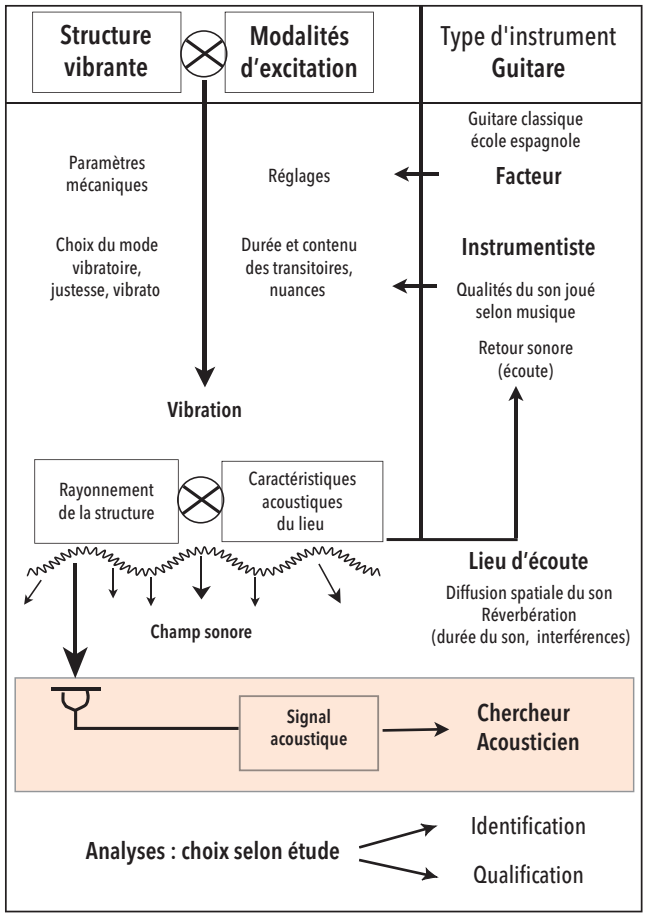


Figure 7.63 Pour évaluer les qualités sonores d'un instrument de musique à partir d'une analyse acoustique, il faut pouvoir s'affranchir de plusieurs variables : celles qui proviennent du champ sonore, celles qui sont dues au jeu d'un musicien et, enfin, disposer d'une typologie acoustique de l'instrument à examiner pour estimer les écarts de qualité de celui-ci par rapport au modèle habituel.

Le facteur construit un modèle donné d'instrument – une flûte traversière à une clé, un pianoforte avec mécanique anglaise, une guitare Ramirez – dont l'architecture et les caractéristiques des matériaux orientent le caractère sonore global de la famille. Ensuite, il règle et ajuste au mieux les différentes parties de l'instrument afin de le doter des potentialités sonores les plus larges possibles.

L'instrumentiste agit sur les paramètres qui contrôlent la production du son. Il est dans bien des cas l'acteur principal des variations de qualités sonores, moyen par lequel il s'exprime musicalement.

L'acousticien capte un signal sonore en un point donné de l'espace. Il doit s'assurer qu'il contrôle les variables supplémentaires que sont l'incidence du lieu d'enregistrement et le rayonnement sonore de l'instrument qui change selon les sons produits.

En fin de compte, le signal sonore qui sera le support des tests d'écoute et à partir duquel sont effectués les analyses acoustiques porte la trace de l'ensemble des actions effectuées à toutes les étapes de la chaîne sonore.

5. Conclusions

La notion de timbre est née en Occident, au milieu du XVIIIe siècle, au sein d'une culture musicale articulée prioritairement sur la hauteur mélodique et harmonique. Elle a été définie par les encyclopédistes comme une qualité du son « musical », c'est-à-dire repérable par sa note fondamentale, sa durée et son intensité, trois grandeurs quantifiables ayant donné lieu à des signes de notation spécifiques. Ainsi défini, le timbre porte en germe les difficultés que nous avons rencontrées pour en faire l'étude acoustique, car le timbre comme qualité du son est plus ou moins implicitement lié à l'instrument qui le produit. D'ailleurs, le terme français « timbre » a fait l'objet de vives discussions au moment de son adoption par les anglophones, certains lui préférant *sound quality*, terme qui correspond à l'Allemand *Klangfarbe* (voir note 53, page 353).

Les encyclopédistes espéraient que seraient rapidement découverts les principes physiques caractérisant le timbre. Deux siècles plus tard, les acousticiens s'en remettent à la sensation auditive, c'est-à-dire à la psychoacoustique, pour définir le timbre. Si certaines caractéristiques du son comme la brillance ou la rugosité sont effectivement des attributs généraux de la perception sonore, liés à des caractéristiques spectrales et temporelles des signaux sonores – comme l'ont montré de nombreux tests acoustiques –, il faut bien admettre aujourd'hui que le concept de timbre est beaucoup plus complexe et qu'il s'agit d'un « objet cognitif » dont l'étude implique la prise en compte des auditeurs et du contexte culturel et musical dans lequel ils s'inscrivent. Ceux-ci, musiciens, compositeurs, instrumentistes et facteurs d'instruments, preneurs de son, fabricants d'enceintes acoustiques, jusqu'aux concepteurs d'automobiles sont, chacun dans leur domaine, des *designers* sonores, des experts remarquables de l'appréciation des qualités du son. La recherche des corrélats acoustiques qui permet d'en rendre compte nécessite donc de définir au préalable ce qui est visé dans ces diverses écoutes, car les stratégies d'orientation perceptive des auditeurs dépendent autant de leurs motivations que de la situation d'écoute et du type de son.

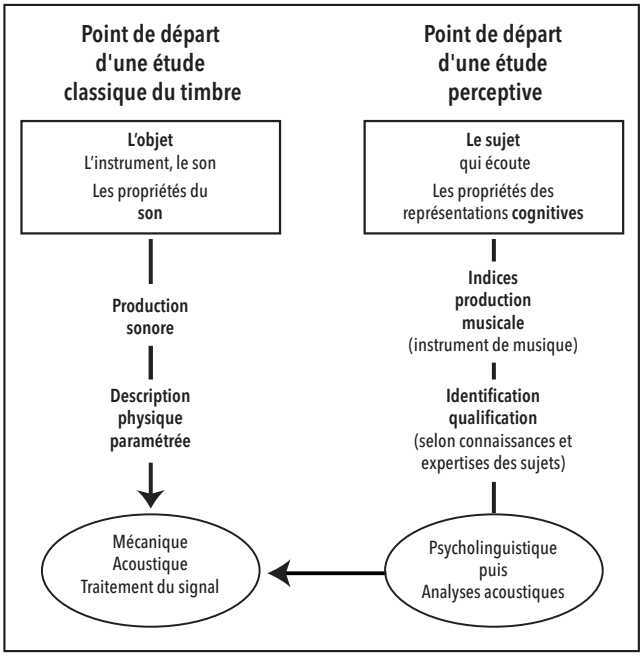


Figure 7.64 Deux approches complémentaires de l'étude du timbre.

D'après Dubois, D., & Castellengo, M., 2006.

Le timbre identitaire

Ainsi posée, l'étude du timbre s'inscrit dans celle plus générale de la catégorisation des sources sonores telle que nous l'avons présentée au chapitre 4. Dans son premier aspect, qui est l'identification des sources, elle repose sur la capacité des auditeurs à repérer dans le signal sonore les régularités acoustiques propres à chaque type de production sonore : le timbre identitaire. En musique, cette catégorisation est particulièrement complexe puisqu'un instrument est précisément conçu pour offrir une palette quasi infinie de variations sonores, par lesquelles l'interprète produit des sons vivants et intéressants à écouter, car toujours changeants.

> *The most musically significant thing about sounds, timbral objects, is not that they are recognizable, identifiable, nor that they are multidimensional wholes, individual and various: it is that they exist in time, have a shape in time, exhibit changes during their time course, and still retain their identity.*[71]
>
> Erickson, R., 1975, p. 58.

Pour autant, la catégorisation du timbre identitaire (ou causal)[72] concerne prioritairement la structure physique des sons comme le prouvent les récents résultats obtenus par les programmes de reconnaissance automatique des instruments de musique. Cependant, par comparaison aux algorithmes numériques actuels, les humains possèdent des stratégies plus efficaces et plus complexes, fondées sur une catégorisation prototypique hiérarchisée en mémoire et sur leur capacité à anticiper

71. Du point de vue musical, la chose la plus importante concernant les sons, ces objets de timbre, n'est pas qu'ils soient reconnaissables et identifiables, ni qu'il s'agisse d'entités multidimensionnelles, uniques et variées ; c'est qu'ils existent dans le temps, qu'ils aient une forme temporelle et qu'ils se transforment au cours de leur existence temporelle, tout en conservant leur identité. (Traduction G. Bloch)
72. Sans vouloir (ni pouvoir) imposer de nouveaux termes pour parler du timbre, il faut bien reconnaître qu'il est important de distinguer en recherche le timbre « identitaire » du timbre « qualitatif ». L'emploi de « couleur sonore » a pour inconvénient d'orienter sur le spectre aux dépens de la temporalité.

et à sélectionner les indices pertinents. Un musicien entraîné est capable de repérer et de suivre un instrument dans la masse complexe d'un orchestre !

Une importante partie de ce chapitre a été consacrée à l'étude de la forme spectro-temporelle des principales familles instrumentales. Nous avons montré, à propos des instruments à percussion, que la notion de timbre se construit différemment selon que les sons possèdent ou non une hauteur définie. Le timbre n'acquiert véritablement une autonomie perceptive que lorsque l'instrument permet de modifier le contenu spectral, indépendamment de la hauteur fondamentale. C'est le cas de la corde vibrante pincée ou frappée et de tous les instruments entretenus ; il s'agit alors du « timbre de la note », au sens classique de la psychoacoustique.

Le timbre qualitatif

Le second aspect du timbre, le timbre qualitatif, concerne plus particulièrement les musiciens. Nous avons montré que les qualités du timbre, qui diffèrent selon les classes d'instruments, portent simultanément sur les aspects temporels et spectraux des champs de liberté de la production du son, sous la dépendance de l'instrumentiste : durée de l'attaque, vibrato, arrêt du son, changement du contenu harmonique. Plus ces possibilités sont larges (violon, flûte traversière), plus l'instrument est difficile à maîtriser et plus il est intéressant du point de vue sonore. La voix humaine, instrument de « timbre » par excellence, aurait mérité un chapitre à part entière en raison de la richesse et de la subtilité des sonorités travaillées par les chanteurs ainsi que de la diversité des modalités perceptives qu'elle suscite puisque la voix donne à entendre, avec la musique, une langue et une personne.

Les écoutes du timbre

Écouter le timbre, c'est développer simultanément une « abstraction timbrale » de la famille instrumentale et une hypersensibilité aux changements à court terme. Si nous sommes capables de penser le timbre de « La » flûte, c'est bien parce que nous avons constitué une catégorie dans laquelle nous avons mémorisé, d'une part, les traits acoustiques communs aux différents sons de cet instrument et, d'autre part, ceux par lesquels il se différencie d'autres instruments à vent comme la clarinette ou le hautbois. Nous y avons inclus les transformations du son de cet instrument en fonction des nuances de jeu, des articulations (lié, détaché, louré) ainsi que des effets qui lui sont spécifiques comme le vibrato, les bruits de transitions entre sons, les bruits de clés, trilles et formules typiques, y compris les caractéristiques sonores propres au goût musical d'une époque. La constitution du prototype porte donc autant sur les indices caractéristiques de la production sonore que sur ceux qui constituent le « beau son » propre à une époque donnée.

Dans l'expérience musicale d'un instrumentiste se développent plusieurs types d'écoutes, tantôt centrées sur la hauteur fondamentale et ses qualités (justesse, vibrato), tantôt sur les qualités du timbre conjointement avec les changements d'intensité, les modalités de jeu et le phrasé musical. L'expertise d'écoute du timbre des instruments de musique exige donc un long apprentissage pour faire la part des variabilités de l'écoute événementielle et celle de l'écoute des qualités esthétiques du son. Cette constatation permet aussi de comprendre les difficultés d'une étude objective du timbre puisque les deux aspects que nous venons de mentionner relèvent, de fait, de la conceptualisation des auditeurs. Le cadre cognitif de l'étude de la perception du timbre doit donc intégrer deux conceptualisations, repérables comme différents types de constructions cognitives, mises en évidence par ailleurs dans les autres modalités sensibles (pour les odeurs et les couleurs, voir Danièle Dubois, 2009). La première conceptualisation, globale et catégorielle requiert des capacités d'abstraction pour saisir les invariants d'une forme temporospectrale ; la seconde, des capacités analytiques grâce

auxquelles nous pouvons discriminer les plus infimes variations des qualités sonores, celles qui sont l'essence même de l'écoute musicale.

Nous avons mis l'accent sur la dimension temporelle, encore trop souvent laissée de côté dans l'étude du timbre alors qu'elle est constitutive de la production sonore et de son écoute. Les techniques d'analyse et de traitement du son offrent maintenant la possibilité de travailler à des échelles de temps différentes, d'inverser les sons, de les disséquer en séparant les parties transitoires des parties tenues. De nombreuses expériences sont encore à faire pour valider perceptivement les hypothèses de recherche de la perception du timbre. En fixant le son, l'enregistrement a aussi modifié notre rapport au temps de l'écoute. Dans la situation « acousmatique » où l'attention est portée au son pour lui-même, dégagé de toute référence à une causalité instrumentale, le terme même de « timbre » disparaît : il n'y a plus d'instrument, seulement du « son ».

Qu'il s'agisse d'un instrument de musique ou de son enregistré, la difficulté centrale de l'étude du timbre reste notre impuissance à en parler de façon simple. Musiciens et instrumentistes produisent de longs et riches discours sur les qualités du son en empruntant des termes à toutes les modalités sensorielles. Il faut toute la science des linguistes pour nous aider à démêler, dans le discours des musiciens, ce qui provient de leur émotion musicale, ce qui traduit leurs jugements sur la technique de jeu ou sur l'interprétation, pour cerner ce que l'on peut rapporter au seul fonctionnement de l'instrument afin de l'interpréter en termes acoustiques.

6. Les sons du chapitre 7

Son 7.1 – La même note la_3, jouée successivement sur sept instruments différents déclenche une écoute prioritaire du timbre « causal », c'est-à-dire une tentative de reconnaissance de l'instrument utilisé. On entend successivement : 1/ trompette ; 2/ clarinette ; 3/ flûte traversière ; 4/ violon ; 5/ saxophone ; 6/ hautbois ; 7/ voix chantée. Sons de la base McGill. [M. C.]

Son 7.2 – Phrase musicale jouée sur la flûte japonaise *shakuhachi*. Le deuxième son subit d'importantes transformations de sonorité qui orientent l'attention de l'auditeur sur les changements de qualité sonore. [M. C.]

Son 7.3 – Musique de percussion, *pakhavaj*. [Voir chapitre 4, Son 4.40]

Son 7.4 – Quelques exemples d'instruments de percussion à hauteur indéterminée regroupés par matière. Bois : *temple block*, *wood bloc*k, clave. Membranes : grosse caisse, tumba, tom-tom. Métal : cloche à vache, triangle, cymbale. Base McGill. [M. C.]

Son 7.5 – Modes de jeu d'une cymbale 16 pouces, mince. Successivement : frappe (*hit*) au centre de la calotte ; frappe au milieu ; frappe au bord ; crash nuance *mf* ; attaque en glissé *mf* avec la brosse ; roulement *crescendo* avec baguettes douces. Base McGill. [M. C.]

Son 7.6 – Instruments de texture sonore : entrechocs de grelots ; entrechocs de tubes de bambou ; roulement de castagnettes ; secouement de maracas *skin* ; secouement de maracas *brass* ; raclement d'une crécelle (*ratchet*). La figure ne montre qu'un extrait de 0,5 s de texture sonore de ces différents instruments. Base McGill. [M. C.]

Son 7.7 – Petite séquence de sons d'instruments à percussion de hauteur indéterminée, réalisée à partir de divers sons de base de données. On entend la grosse caisse, le triangle, la *cabasa*, le tambour de basque, la caisse claire (sans timbre),

les *wind chime*s, le *wood block,* le tambour militaire (roulement). [Voir chapitre 2, figure 2.40]

Son 7.8 – Trois sons de percussion donnant la même note sol_2 ; timbale, marimba, vibraphone. Base McGill. Si vous entendez la timbale une octave plus bas que le marimba et le vibraphone, allez au chapitre 6, § 3.4.3. [M. C.]

Son 7.9 – Six sons de percussion donnant la même note sol_3 : xylophone, marimba, célesta, vibraphone, cloche tube, vibraphone joué à l'archet. base McGill. [M. C.]

Son 7.10 – Quatre variations sur deux instruments et les octaves : xylophone et marimba. Dans certains enchaînements, le partiel 2 peut devenir dominant en modifiant l'intervalle perçu ou en donnant naissance à une deuxième ligne mélodique. Dans l'exemple 4, certains auditeurs entendent une quinte descendante. Base McGill. [M. C.]

Son 7.11 – La même note la_3 jouée sur des instruments à percussion accordés et sur des instruments à cordes libres. Successivement : vibraphone, marimba, célesta ; puis piano, clavecin, guitare. Son de piano : base RWC 011NOM ; les autres, base McGill. [M. C.]

Son 7.12 – La même « note » de guitare, mi_2, jouée de six façons différentes. E. Pélissier. [V. Mons]

Son 7.13 – Un son instrumental bref la_3, est donné à entendre trois fois de suite. Quel instrument l'a produit ? Voir texte. [M. C.]

Son 7.14 – Série des sons instrumentaux, note do_3 de la figure 7.15. Première série, silence de 3 s, puis deuxième série. Dans l'ordre, sons issus des instruments suivants : violon, flûte, hautbois, clarinette, saxophone, accordéon, trompette. Bases de données sources indiquées sur la figure. [M. C.]

Son 7.15 – Série des sons instrumentaux note la_3 de la figure 7.16. Dans ce montage, on entend deux sons différents pour chaque instrument. Même ordre des sources sonores que pour le do_3 ; violon, flûte, hautbois, clarinette, saxophone, accordéon, trompette. Bases de données sources indiquées sur la figure. [M. C.]

Son 7.16 – Les deux sons de trompette et de flûte dont le transitoire est analysé. Base Iowa. [M. C.]

Son 7.17 – Jeu de guimbarde dans lequel le musicien joue un son sans souffle dont la sonorité surprend. On l'entend clairement à la septième battue (en comptant 1 sur le premier temps). Le spectre impair paraît dissonant : comparez avec la dernière note. Disque 33 t. *Le Chant du Monde,* LDX 74434, coll. Spécial instrumental *La Guimbarde.* [John Wright]

Son 7.18 – Trompette : transformations spectrales par les sourdines. La note do_3, trompette normale puis avec trois sourdines différentes : bol, sèche, wah-wah. Base RWC (212TRNOF ; 212TRC1F ; 212TRM7F ; 212TRW1F. [M. C.]

Son 7.19 – Le son original (la_3) de violon utilisé dans le montage « chimère » (son 7.13), joué une fois. [M. C.]

Son 7.20 – Mélodie jouée sur un instrument coréen, le *taegum* qui est une flûte traversière de grosse taille, munie d'un mirliton. La richesse harmonique et les attaques extrêmement précises font plutôt penser à un instrument de la famille du hautbois. [Archives LAM]

Son 7.21 – Jeu d'orgue dit jeu de tierce. Orgue de la cathédrale de Poitiers avant restauration, 1986. [Archives LAM]

Son 7.22 – Voix « hélium ». Incidence de la transformation des fréquences des formants indépendamment de la fréquence fondamentale de la voix. Phrase parlée : *This is my voice*. Tout d'abord voix modifiée puis voix normale. Ensuite la voyelle « a » modifiée et normale. La hauteur fondamentale de la voix reste inchangée et pourtant la voix semble plus haute lorsqu'elle est affectée par l'hélium. [Joe Wolfe]

Son 7.23 – Gammes chromatiques rapides ascendantes (quelques-unes avec descente) jouées *legato*. Son 7.23a : hautbois, 5 s, (L. Debray). Son 7.23b : clarinette, 8 s, (Y. Ettlinger). Son 7.23c : saxophone, gamme ascendante descendante, 8 s, (C. Delangle). Son 7.23d : basson, gamme ascendante descendante, 9 s, (M. Allard). Son 7.23e : flûte traversière, 5 s, (B. Fabre). Son 7.23f : trompette, 5 s, (L. Roquin). Son 7.23g : violon, 7 s, (A. Gill). Son 7.23h : accordéon, 8 s, (P. Monichon). [Archives LAM]

Son 7.24 – Gammes chromatiques rapides ascendantes en jeu détaché. Son 7.24a : flûte traversière (P.-Y. Artaud). Son 7.24b : saxophone (ascendant/descendant (C. Delangle). Son 7.24c : trompette (L. Roquin). Son 7.24d : violon (A. Gill). [Archives LAM]

Son 7.25 – Des sons issus d'instruments différents auxquels on affecte la même enveloppe temporelle semblent provenir de la même source acoustique. Exemple 7.25a : huit sons possédant la même enveloppe dynamique. Exemple 7.25b : les huit sons originaux. P. Schaeffer, *Solfège de l'objet sonore*, CD 2, index 28 et 29, 1967-1998. [Ina-GRM]

Son 7.26 – Un bien curieux son vibré qui est un authentique son de piano devenu méconnaissable, par synthèse croisée avec l'amplitude d'un son de flûte. Son 7.26a : son transformé ; Son 7.26b : son original. [M. C.]

Son 7.27 – Variabilité du timbre identitaire : rôle de l'interprète. Sons isolés en arpège extraits de bases de données. Son 7.27a : violoniste 1. Son 7.27b : violoniste 2. Son 7.27c : flûtiste 1. Son 7.27d : flûtiste 2. Base McGill (a, d) ; RWC152NOM et RWC331VIM (b, c). [M. C.]

Son 7.28 – Exemple de diversité des attaques au saxophone alto. Note *ré*$_3$ (hauteur réelle). Articulation « normale » (*de*) ; plus nette (*te*) ; attaque dure (sons séparés) ; sons dit « en cloche » ; son en cloche inversée. Les deux derniers exemples sont aspirés. C. Delangle. [Archives LAM]

Son 7.29 – Six notes isolées de flûte, *la*$_3$, extraites des bases de données. Successivement : flûte traversière ; flûte à bec baroque ; flûte de pan ; et *shakuhachi* selon trois modes de jeu. Le timbre identitaire *flûte* se décline selon une infinité de sonorités remarquablement adaptées aux différents types de musique. Lorsqu'on écoute une seule note – ce qui permet de s'abstraire des caractéristiques dues au style de musique – on remarque immédiatement les points communs à ces cinq sons (la prédominance du fondamental et la présence de souffle), et en même temps les énormes différences de qualité sonore selon qu'il s'agit de musiques de notes ou de musiques de sonorité. Base McGill (Sons 1 et 2) ; base RWC : son3, 333VIM ; son 4 : 351SHMRF ; son 5 : 351351SHMIF ; son 6 : 351SHKAF. [M. C.]

Son 7.30 – Modes de jeu contemporains à la flûte traversière. Successivement : 1-*whistle tones* ; 2-son de « trompette » ; 3-sons éoliens jeu lié puis détaché ; 4-jet whistles ; 5-percussions de clés ; 6-pizzicati fort et détaché puis légers ; 7-*tongue-ram*. P.-Y. Artaud ; 1977. [Archives LAM]

Son 7.31 – Mélodie « arlequin » réalisée par montage d'instruments différents enregistrés au LAM : flûte, piano, clavecin, voix. [M. C.]

Son 7.32 – Écoute de la même mélodie jouée avec trois instruments à cordes libres. Son 7.32a : clavecin (M.C.). Son 7.32b : guitare (E. Pélissier) : Son 7.32c : piano (M. C [M. C.]

Son 7.33 – Écoute de la même mélodie jouée par trois instruments à vent. Son 7.33a : flûte traversière (B. Fabre). Son 7.33b : clarinette (J.-M. Dussert). Son 7.33c : trombone à coulisse (B. Sluchin). [Archives LAM]

Son 7.34 – Écoute de la même mélodie vocalisée sur « a » puis chantée avec un texte (M. Garnier). [M. C.]

Son 7.35 – Une note de flûte traversière ($ré_4$), jouée quatre fois par le même musicien. Première note de la phrase musicale du Son 7.36. Dans l'ordre : flûte basse, flûte en ut, flûte alto, piccolo. [M. C.]

Son 7.36 – Famille de flûtes. Une mélodie jouée à la même hauteur sur les quatre flûtes : piccolo, grande flûte, flûte alto, flûte basse. P.Y. Artaud, 1985. [M. C.]

Son 7.37 – Famille de flûtes. Montage d'écoute de la dernière note de chaque mélodie. Dans quel ordre sont joués les instruments ? Réponse : voir note de bas de page 389. P.-Y. Artaud. [M. C.]

Son 7.38 – Famille d'instruments du quatuor. Une mélodie jouée à la même hauteur fondamentale sur chacun des instruments. Son 7.38a : le violon. Son 7.38b : le violon alto. Son 7.38c : le violoncelle. A. Laracine, cours de prise de son à la Fémis, 1981. [M. C.]

Son 7.39 – Un guitariste et une guitare (baroque). Enregistrement A, musique de François Le Cocq : *Pièces pour Guitare* (1729). Musique en Wallonie, MW 80045. [Rafaël Andia]

Son 7.40 – Même guitariste et même guitare avec un autre preneur de son. Enregistrement B, musique de Robert de Visée : *L'oeuvre pour Guitare,* Harmonia Mundi-CNRS, HM 60x3. [Rafaël Andia]

Son 7.41 – Quatre paires de sons de l'expérience de J. M. Grey. Note mib_3. Son 7.41a : paire 2-12. Son 7.41b : paire 2-1. Son 7.41c : paire 10-2. Son 7.41d : paire 13-12. Thèse O. Hourdin, 1995. [M. C.]

Son 7.42 – Expérience de J. M. Grey. Classement des sons dans l'ordre de l'axe X (dimension 1). Successivement : 13, 2, 1, 7, 3, 5, 14, 9, 10, 11, 8, 15, 4, 6, 16, 12. [*Ibid.*]

Son 7.43 – Expérience de J. M. Grey. Classement des sons dans l'ordre de l'axe Y (dimension 2). Successivement : 10, 14, 15, 16, 13, 11, 4, 12, 2, 1, 3, 6, 8, 5, 9, 7. [*Ibid.*]

Son 7.44 – Expérience de J. M. Grey. Classement des sons dans l'ordre de l'axe Z (dimension 3). Successivement : 6, 5, 16, 7, 14, 15, 10, 8, 1, 9, 2, 12, 3, 13, 4, 11. [*Ibid.*]

Son 7.45 – Les 13 sons du test steel-drum ; note $ré_2$ (146 Hz). [P. Gaillard, 2000 et 2007]

Son 7.46 – Quelques sons du test piano ; note do_1 (62 Hz). Son 7.46a : 2A. Son 7.46b : 2D. Son 7.46c : 5D. Son 7.46d : 5A. Son 7.46e : 1B (son orignal, non modifié). Voir texte pour le détail des paramètres. [J. Bensa, 2005]

Son 7.47 – Séquence musicale à écouter crayon en mains avec pour consigne : « vous allez entendre une séquence musicale ; qualifiez le son de l'instrument ». [M. C.]

Son 7.48 – Clarinettiste jouant une phrase musicale avec deux qualités de timbre : son « détimbré » puis son « timbré ». [LAM]

Son 7.49 – Flûtiste jouant une phrase musicale avec deux qualités de timbre : son « détimbré » puis son « timbré ». [LAM]

Son 7.50 – Un baryton chante la phrase « *Il vole là-haut tout en oubliant nos âmes* ». Voix «détimbrée» ; voix «timbrée». D. Sotiropoulos, 2004. [LAM]]

Son 7.51 – Séquence de voyelles chuchotées donnant à entendre une mélodie connue. Voix féminine. [M. C.]

CHAPITRE 8

SYSTÈMES D'INTERVALLES ET ACCORDAGE

1. La notion d'intervalle

1.1. Données perceptives

1.1.1. Les seuils différentiels et les intervalles musicaux

Un intervalle est un écart, une distance entre deux éléments conceptuellement similaires. La capacité à apprécier un intervalle entre deux sons est une donnée fondamentale de la perception que nous avons déjà abordée à propos des seuils différentiels, les plus petits intervalles perceptibles (chapitre 3 § 2.5). Ces seuils différentiels de fréquence et d'intensité rendent compte de la sensibilité perceptive au changement que peut subir un son donné dans un bref espace de temps. Plusieurs chercheurs ont soulevé l'apparente contradiction qui existe entre la faible valeur des seuils différentiels témoignant du pouvoir de discrimination de l'oreille, et celle, cinq à dix fois plus grande, des intervalles en usage dans la musique[1]. La raison en est que les intervalles qui structurent une musique donnée – intervalles de durée pour le tempo et pour les rythmes, intervalles de hauteur pour la mélodie et l'harmonie[2] – sont des grandeurs cognitives qui relèvent de la catégorisation perceptive et non de la seule discrimination sensorielle. Comme pour les couleurs ou les sons de la langue, les catégories d'intervalles de hauteur sonore sont le fruit d'élaborations individuelles au sein d'une culture donnée, celle pour laquelle ils font sens, même si, selon certains auteurs, le fait d'être « imprégné » d'une culture musicale donnée dès la plus petite enfance peut créer l'illusion d'une connaissance innée. Apprendre la musique c'est, entre autres, apprendre à écouter les intervalles afin de les reconnaître, voire les nommer. C'est aussi apprendre à faire abstraction de la variabilité des intervalles dans la musique vivante, grâce à la perception catégorielle.

1.1.2. La perception catégorielle et la tolérance perceptive

Par perception catégorielle, on entend le fait que les objets et les événements du monde, avec lesquels nous interagissons, font partie d'un réseau organisé en catégories distinctes[3] propres à une culture. Les individus d'une société donnée partagent un certain nombre de catégories dont relèvent les sons de la langue, les couleurs et les intervalles musicaux. Le fait que des mots soient associés aux caté-

1. Par exemple le seuil différentiel en fréquence d'un son de 500 Hz à 60 dB (6,5 cents) est 15 fois plus petit que le plus petit intervalle de la musique européenne : le demi-ton (100 cents). Voir chapitre 3, figure 3.20.
2. Les tentatives de systématisation d'intervalles de timbre n'ont pas abouti en raison de la complexité perceptive de cette notion. Voir chapitre 7.
3. La catégorisation est présentée au chapitre 4, § 4.

gories et aux éléments qu'elles renferment renforce la discrimination et la reconnaissance de ces éléments.

En France, l'apprentissage musical est associé aux sept degrés de la gamme diatonique (*do, ré, mi, fa, sol, la si*) et à leurs altérations par les dièses et les bémols. Le plus petit intervalle de ce système est le demi-ton qui, selon l'accord en tempérament égal (voir page 431), est la douzième partie de l'octave. Écoutons le Son **8.1a** dans lequel un intervalle de seconde majeure mélodique *(do-ré)* se répète plusieurs fois en s'agrandissant. Un musicien occidental, sensible aux modifications de l'intervalle, perçoit un ton qui s'élargit progressivement et qui, soudain, passe sans transition du ton trop grand à une tierce trop petite, car pour lui il n'existe pas d'intervalle intermédiaire entre le ton et la tierce mineure. Un tel comportement est caractéristique de la perception catégorielle[4]. Des observations similaires ont été faites pour la perception des sons de la langue, en particulier pour les voyelles. À l'intérieur d'une catégorie donnée, nous pouvons donc accepter des variations plus ou moins importantes autour de la valeur théorique de l'intervalle, ce que l'on désigne par « tolérance perceptive »[5]. Les instruments mélodiques, par exemple, permettent de produire une infinie variété de grandeurs du même intervalle formel pour répondre à l'attraction mélodique des notes « sensibles » proches de la tonique ou de la dominante.

Interpolation entre deux intervalles

Son 8.1a (42″)

Notez le numéro pour lequel vous basculez à l'intervalle supérieur (voir page 439).

Son 8.1b (42″)

Même test dans le sens décroissant

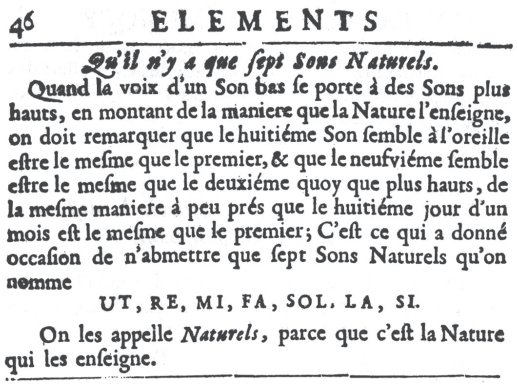
Les Sons **8.1a** et **8.1b** font prendre conscience des difficultés auxquelles on se heurte pour traiter des intervalles des échelles de hauteur. Ils doivent être définis avec suffisamment de précision pour être différenciables, alors que leur réalisation musicale donne lieu à une variabilité souvent importante, que seuls les initiés sont capables d'apprécier. Cet état de fait impose de « naviguer » entre théorie et pratique musicale, entre mesures et écoute, et c'est ainsi que nous aborderons le thème des intervalles musicaux.

1.2. Données musicologiques

Note préliminaire : dans la communauté des musicologues, la notion d'intervalle appelle celles d'échelle, de gamme, de mode, de système intervallique, mais le sens de ces termes peut changer selon les pratiques musicales. L'étude qui suit s'attache prioritairement à la caractérisation acoustique de l'écart entre deux sons, au sens de degrés d'une échelle, sans a priori sur leur fonction musicale.

Nous vivons, en Occident, au sein d'une société où la musique, omniprésente dans les lieux publics et les médias (radio, télévision), paraît naturelle. Elle est enseignée à l'école et on peut l'apprendre via les innombrables méthodes disponibles sur Internet : « *do, ré, mi, fa, sol, la, si, do* », chante-t-on sans se poser de question[6]. Mais d'où viennent ces notes ? En quoi sont-elles différentes de celles que jouent les musiciens orientaux ou africains, dont la musique sonne bizarrement à nos oreilles ? Le choix des intervalles et des degrés de la gamme est-il pure convention ou existe-t-il des principes communs aux différentes musiques du monde ? S'agit-il d'une donnée de la nature, comme on pouvait le lire dans ce traité du XVIIe siècle (figure 8.1) ? L'acoustique est souvent sollicitée pour fournir des réponses à ces questions qui préoccupent les musicologues.

4. Burns, 1999, p. 226.
5. Voir Leipp, 2011, p. 132 ; Francès, 1984, p. 23 pour « l'abstraction notale ».
6. Ce à quoi le lecteur averti répondra, en chantant par mouvement descendant : « gratt' moi la puce que j'ai dans l'dos ».



46 ELEMENTS

Qu'il n'y a que ſept Sons Naturels.

Quand la voix d'un Son bas ſe porte à des Sons plus hauts, en montant de la maniere que la Nature l'enſeigne, on doit remarquer que le huitiéme Son ſemble à l'oreille eſtre le meſme que le premier, & que le neufviéme ſemble eſtre le meſme que le deuxiéme quoy que plus hauts, de la meſme maniere à peu prés que le huitiéme jour d'un mois eſt le meſme que le premier; C'eſt ce qui a donné occaſion de n'abmettre que ſept Sons Naturels qu'on nomme

UT, RE, MI, FA, SOL, LA, SI.

On les appelle *Naturels*, parce que c'eſt la Nature qui les enſeigne.

Figure 8.1 Les sept *Sons Naturels* de la gamme diatonique en 1696.

Loulié, E., 1696, p. 46 [Minkoff].

1.2.1. Les intervalles en musique : mélodie, contrepoint, harmonie

La musicologie occidentale dispose d'un considérable corpus d'ouvrages théoriques dont la filiation remonte à l'Antiquité grecque. Bien que la réalité sonore de la musique effectivement jouée à l'époque de Platon nous soit à tout jamais inaccessible, ces ouvrages nous ont transmis les principes théoriques élémentaires qui ont fondé l'organisation des intervalles entre les sons de l'échelle musicale. Copiés, traduits plusieurs fois au cours des siècles, les écrits antiques ont suscité un nouvel intérêt autour du Xe siècle, époque où est apparue la notation de la musique[7]. Les documents les plus anciens, retrouvés dans les abbayes, nous ont transmis des mélodies chantées. Pour la question qui nous concerne, celle des intervalles, deux changements majeurs se produisent au cours des siècles. Dès le XIe siècle se développe le contrepoint – un chant à plusieurs voix simultanées – qui implique de formaliser, en plus des intervalles entre les sons successifs d'une mélodie, les intervalles entre sons simultanés se produisant à la rencontre de deux voix différentes. Ce sont l'octave, la quinte et la quarte. La polyphonie se développe aux XVe et XVIe siècle, ainsi que la pratique d'instruments de hauteur fixe à clavier, tels l'orgue et le clavecin, lesquels nécessitent de définir précisément les grandeurs des intervalles entre les sons. Durant cette période riche en discussions et polémiques sur le nombre de sons dans l'octave, les traités fournissent la description de plusieurs systèmes d'accordage qui s'opposent principalement sur le statut de l'intervalle de tierce majeure dans l'accord de trois sons dit accord parfait (ex. *do, mi, sol*). Au XVIIIe siècle, l'harmonie introduit le concept de « basse fondamentale » et modifie les relations entre les sons de la mélodie et ceux des accords sous-jacents, au point que pour certains auteurs la mélodie naît de l'harmonie. Les nombreuses transformations qui se produisent au XIXe siècle imposent progressivement un schéma intervallique « passe-partout », connu de longue date : la division de l'octave en douze demi-tons égaux. Ce système permet les modulations de la nouvelle musique ; il s'accommode du « flou » de hauteur de l'orchestre, accentué par la pratique du vibrato[8], et convient particulièrement bien à l'accordage du piano dont l'expansion est rapide (voir infra § 3.3.2 ; voir aussi chapitre 6, § 3.2).

Dans la plupart des autres cultures, la musique s'est développée et complexifiée tout en restant fondamentalement mélodique. Elle présente de grands raffinements

7. Voir Olivier Culin, 2006, *L'image musique*.
8. Voir chapitre 6, § 4.3, le Son 6.54, pour un exemple de cette tolérance.

dans le statut des degrés de l'échelle musicale : multiplication du nombre de degrés disponibles dans l'octave, grandeur des intervalles, élaboration de « modes » : systèmes musicaux complexes combinant non seulement les degrés de l'échelle mais aussi certains schémas rythmiques et des formules d'ornementation spécifiques. Chaque mode convient à une expression musicale particulière[9].

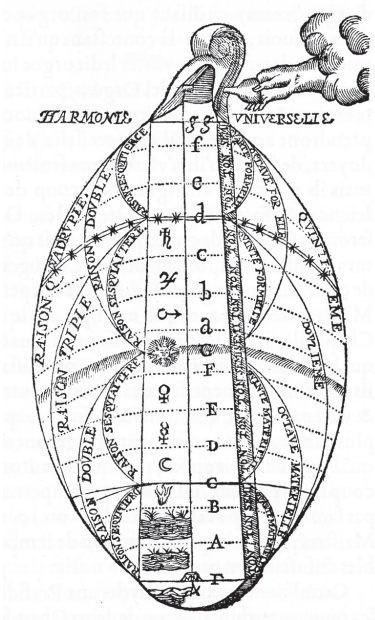


Figure 8.2 Grande lyre de l'univers de Mersenne (1636), reprise de Robert Fludd (1574-1637).

Mersenne, M., 1636, *Livre de l'utilité de l'harmonie universelle et des autres parties des mathématiques*, Proposition X,, p. 49 [tome 3 du fac-similé CNRS, 1965].

De ce très bref et schématique survol nous retiendrons que les deux situations, *monodie* et *polyphonie,* conduisent à des orientations différentes quant à la grandeur des intervalles, et doivent donc être étudiées séparément. Remarquons aussi que les traités théoriques tendent, dans toutes les cultures, à schématiser la réalité sonore. La matière enseignée sous le nom de *musica* pendant le Moyen Âge était une science des rapports numériques[10], assez éloignée de la pratique musicale de l'époque désignée par *cantus*. Par ailleurs, les théories musicales exposées dans la plupart des traités mêlent étroitement les considérations numériques, instrumentales et symboliques, car leur auteur vise à élaborer une harmonie générale, une cosmogonie universelle. Cette vue globale de la théorie des intervalles musicaux persiste au XVIIe siècle dans l'ouvrage de Marin Mersenne[11] (voir figure 8.2).

Parallèlement aux musiques savantes théorisées et enseignées dans les institutions, coexistent des musiques dites de tradition orale, toujours vivaces aujourd'hui. La plupart d'entre elles sont d'essence mélodique, et rares sont celles qui sont constituées uniquement de sons continûment variables. Le glissé, l'indécis dans l'intonation relèvent de la parole, et il semble que, dans toute musique mélodique, il existe au moins un son stable de référence, autour duquel s'organise un système de degrés plus ou moins bien définis. Le Son **8.2** offre un intéressant exemple de musique organisée de part et d'autre d'un degré médian[12]. Si nous posons que toute musique mélodique implique l'existence d'un cadre intervallique donné, les questions que se pose le chercheur sont celles du nombre, de la stabilité et de l'importance relative de ces intervalles et, en particulier, celle des principes qui président aux choix des intervalles observés. La plupart des auteurs (Brailoiu, 1953 ; Chailley, 1954-1955 ; Labussière, 2006 et 2007, Fürniss, 2000) admettent qu'il existe au moins deux repères (ou piliers) communs à la quasi-totalité des systèmes musicaux, qui sont : l'intervalle de *quarte* comme cadre structurant les systèmes monodiques, et les intervalles *d'octave* et de *quinte*

Son 8.2 (26")

Chant mixte hommes, femmes, enfants (Érythrée)

9. À titre d'exemple, le lecteur pourra consulter Mukherjee (2004) pour l'Inde ; Beyhom (2010) pour le monde arabo-persan ; Trébinjac (2008) pour l'Extrême-Orient (Chine) ; Arom *& coll.* (1997, 2000) pour l'Afrique.
10. Avec l'arithmétique, la géométrie et l'astronomie, elle faisait partie du *quadrivium*, les quatre sciences se rapportant aux mathématiques.
11. « La grande Lyre de l'univers, que le divin Orphée gouverne en donnant tel ton et tel accord qu'il lui plaist à toutes les parties du monde comme l'on peut comprendre par ceste figure, dans laquelle les lettres ordinaires de l'échelle de Musique... représentent chaque estage du monde » *op. cit.*
12. En l'occurrence un *sib*. La mélodie a comme autre repère le *fa* à la quinte supérieure et, implicitement son octave inférieure, présente par l'union des voix d'hommes et de femmes en alternance, comme il en est fait mention dans le paragraphe suivant. Entre ces deux repères, la voix « glisse ».

(ou *quarte*) pour les systèmes de sons simultanés. Cette observation trouve une explication aisée en acoustique.

1.2.2. Les musiques de tradition orale chantées

La voix tient une place singulière dans les cultures du monde entier. C'est la voix qui porte la parole ; elle est à la disposition de tout individu ; enfin, elle est un « instrument » de musique dont le son est riche en harmoniques. Les premiers intervalles de la série harmonique, toujours présents dans le son vocal, sont facilement isolables par l'oreille[13] et, du fait de la mobilité des résonateurs buccaux, certains peuvent être amplifiés comme nous l'avons vu à propos du chant dit harmonique (voir chapitre 6, § 2.7.2). Mais les voix de femmes, d'hommes et d'enfants évoluent dans des tessitures différentes : pour chanter ensemble, « d'une seule voix », ils se retrouvent le plus souvent à intervalle d'octave, plus rarement en quintes (ou quartes) parallèles. Écoutez à nouveau le Son **8.2**. Les membres de ce groupe ne cherchent pas à produire une polyphonie : ils chantent « la même chose », avec l'aisance que chacun éprouve en restant dans sa tessiture, tout en ressentant fortement la fusion sonore produite par les nombreuses coïncidences harmoniques entre les voix[14]. La similitude : consonance d'octave, voire de quinte, est un fait d'expérience courante, dès qu'un enfant chante avec un adulte. Il en va autrement pour les intervalles entre les degrés d'une mélodie. Le confort vocal conduit plutôt à éviter de grands intervalles, comme l'octave, difficiles à entonner. L'ambitus de la plupart des mélodies dépasse rarement la sixte.

Dans la musicologie occidentale, l'expression numérique des intervalles d'octave, de quarte et de quinte correspond aux premières divisions du monocorde.

1.2.3. Du monocorde aux harmoniques effleurés d'une corde

> *...muni de sa harpe, l'homme peut aisément témoigner de préoccupations nouvelles : [...] visualiser ce qu'il fait [...] traduire le son en quelque chose d'autre, [...] théoriser. »*
>
> Dampierre, (de) E., 1998, p. 18.

L'usage de la corde vibrante a certainement joué un rôle important dans la systématisation des intervalles musicaux. On peut remarquer que les musiques modales, citées plus haut, se sont développées dans les cultures ayant adopté des instruments à cordes multiples, de facture élaborée, telles les cithares, les harpes et les luths. Quant au monocorde dont il est question dans les traités anciens, c'est à la fois un instrument pédagogique et un instrument de mesure qui a subsisté jusqu'au XXe siècle[15].

Figure 8.3 Le monocorde de Salomon de Caus.

Salomon de Caus, 1615, page de couverture [Minkoff].

13. Voir le chapitre 3, § 2.2, au sujet des bandes critiques et de la séparation des composantes spectrales.
14. Le chant en octave est cité par Chailley, 1950, *Histoire musicale du Moyen Âg*e, p. 71, avec une référence à Aristote.
15. C'est avec un monocorde de grandes dimensions muni d'une corde fine et bien calibrée que les scientifiques du XIXe siècle, doués d'une bonne oreille, pouvaient faire des mesures précises de fréquence avant le développement de l'électronique au XXe siècle.

Lorsqu'on dispose d'une corde longue et souple, fixée aux deux extrémités d'une caisse rectangulaire et d'un chevalet mobile que l'on peut déplacer continûment pour l'arrêter avec précision à une position donnée, il est possible de comparer à l'oreille les sons émis par les deux portions de la corde, et de mesurer les rapports de longueurs de ces deux portions (voir les figures 8.3 et 8.4).

Figure 8.4 Démonstration sur un monocorde, des proportions numériques entre les deux sons d'un intervalle.
Fogliano, 1529, *Musica theorica*, feuille XIII. [BMVR de Nice - Bibliothèque Romain-Gary.]

Dès le XIIe siècle, les traités fournissent des schémas comportant les noms grecs puis latins des intervalles, noms qui renvoient explicitement au rapport numérique qui permet de les obtenir en modifiant la longueur de la corde : *diapason* ou *dupla* (double) pour l'octave ; *diapente* ou *sesquialtera* (une fois et demie) pour la quinte ; *diatessaron* ou *sesquitertia* (quatre tiers) pour la quarte. À la suite de Marin Mersenne (1636), qui lie la production du son à la vibration plus ou moins rapide de la corde, Joseph Sauveur (*Principes d'acoustique et de musique*, 1701) énonce la correspondance, inverse, entre les longueurs de corde et les fréquences de deux sons formant un intervalle donné. Il définit une unité de mesure logarithmique des intervalles[16] propre à représenter ceux d'un système quelconque. Sauveur fait aussi l'expérience de sons qu'il nomme harmoniques (voir page 51 de son ouvrage), ceux

16. Le *Méride*, 43e partie de l'octave et *l'Eptaméride*, 301e partie de l'octave qui, curieusement, prendra ensuite le nom de *savart* et non de *sauveur*.

qu'il obtient en effleurant la corde[17]. En observant le nombre des ondulations avec leurs *nœuds* et leurs *ventres*, ainsi que les intervalles que font ces sons harmoniques avec le son de la corde entière, il dresse un tableau dans lequel les intervalles et leur mesure logarithmique apparaissent dans l'ordre croissant du nombre de vibrations rapporté au son fondamental, et non plus dans l'ordre de l'échelle diatonique. Même si l'ouvrage de Sauveur a découragé les musiciens de son époque, tant par le caractère théorique de son contenu que par les nouveautés qu'il impose dans la dénomination et l'écriture des sons, le terrain est prêt pour que naisse la théorie de l'harmonie que Rameau publie à Paris en 1722[18].

1.2.4. La série harmonique en théorie musicale

La série harmonique a une double vie musicale. Elle a initialement été présentée comme une suite immuable de sons et d'intervalles, dont l'importance pour la théorie musicale était inversement proportionnelle à leur rang d'apparition. Les premiers intervalles, dont le rapport s'exprime aussi par les premiers nombres : 2 pour l'octave ; 3/2 pour la quinte ; 4/3 pour la quarte ; 5/4 pour la tierce majeure et 6/5 pour la tierce mineure, étaient définis comme *consonants*. Les autres : ton majeur 9/8 ; ton mineur 10/9 ; demi-ton majeur 16/15 ; demi-ton mineur 25/24, étant *dissonants*. On aura remarqué que les auteurs s'arrêtent à l'harmonique 6, « le nombre 7 ne pouvant donner aucun intervalle agréable », écrit Rameau en 1722 (figure 8.5).

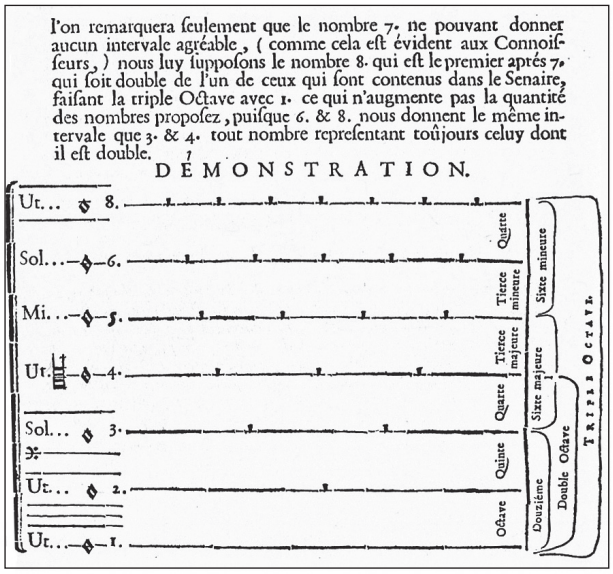


Figure 8.5 Comme le précise ici Jean-Philippe Rameau pour les « Connoisseurs », les nombres pairs, qui sont des octaves de sons inférieurs, n'entrent pas dans le décompte des harmoniques.

Rameau, J.-Ph., 1722, extrait de la page 4 [Bnf-Gallica.]

17. Rappelons que les harmoniques effleurés sont en fait des modes propres de la corde. Voir chapitre 1, § 2.4.4.
18. Pour un exposé détaillé des rapports entre théorie musicale et connaissances scientifiques, voir la remarquable thèse de Théodora Psychoyou, 2003.

Les discussions sur l'intégration de l'harmonique 7, et d'autres de rang supérieur, reprendront avec les transformations de l'écriture musicale aux XIXe et XXe siècles (Fichet, 1996). Mais lorsque paraissent les appareils d'analyse et de filtrage du son et que les musiciens s'emparent de la synthèse sonore, la série harmonique devient alors un « objet » musical à part entière qui renouvelle l'écoute des hauteurs. La hiérarchie décroissante établie selon le rang de l'harmonique s'abolit, puisqu'il devient possible d'amplifier n'importe lequel des harmoniques.

1.3. Les échelles mélodiques de sélection d'harmoniques : le problème de l'harmonique 7

Parallèlement au long processus qui aboutit, en Occident, à l'intégration des intervalles et des accords de la musique dans une théorie générale de l'échelle harmonique qui trouve sa justification au XIXe siècle, particulièrement avec les travaux de Helmholtz, de nombreuses cultures ont développé des musiques fondées dès l'origine sur l'échelle intervallique des harmoniques.

Nous présentons trois tech-niques de jeu se fondant sur le même principe acoustique : le couplage des résonateurs buccaux avec une source sonore stable, riche en harmoniques. Cette source peut être la corde d'un arc musical, la languette vibrante d'une guimbarde, ou simplement les « cordes vocales » du chanteur[19]. La mélodie produite naît de la sélection d'un ou plusieurs harmoniques du spectre de la source dont la hauteur fondamentale, toujours présente, sert d'assise.

1.3.1. Le chant harmonique ou diphonique

Son 8.3 (29")

Chant harmonique (Mongolie)

L'exemple musical **8.3** est emprunté à la tradition mongole. Passé l'étonnement provoqué par ce type de chant chez un auditeur qui l'entend pour la première fois, la mélodie est aisée à écouter et à reproduire[20]. Un musicien peut la noter comme indiqué sur la figure 8.6-A, en précisant qu'il s'agit d'un air pentatonique sur les notes « *ré, mi, sol, la, si* ». Examinons l'analyse spectrale de ce chant sur le sonagramme de la figure (8.7).

Son 8.4 (16")

Début du chant harmonique, deux fois plus lent

La maîtrise du chanteur est telle que chaque harmonique de la mélodie est sélectionné avec une grande netteté. Il est aisé de réaliser, à partir du spectre, une transcription acoustique exacte sur une grille harmonique (représentée au-dessous du sonagramme). La hauteur fondamentale de la voix, dont on voit les harmoniques H1 et H2, oscille de part et d'autre d'une note stable, un sol_2 de fréquence 200 Hz. L'ambitus de la mélodie s'étend entre les harmoniques 5 et 13, soit entre 1000 et 2600 Hz (figure 8.9). Les trois premiers sons de la mélodie sont les harmoniques 6, 7, 8. Or, puisque H6 est un *ré*, comment se fait-il que pratiquement tous les auditeurs entendent et notent : *ré, **mi**, sol* et non *ré, **fa**, sol* ? Bien qu'un peu « bas » aux oreilles d'un occidental, l'harmonique 7 d'un *sol* est un *fa*. On peut s'en convaincre en écoutant le début du Son **8.4** ralenti tout en se préparant intérieurement à entendre *ré, **fa**, sol.*

Le 2e et le 4e son de la mélodie sont étrangers à ceux de notre gamme diatonique, mais nous les assimilons aux degrés les plus proches d'une échelle pentatonique familière, et d'autant plus aisément en présence de vibrato (voir chapitre 6, § 4.3.2).

19. En toute rigueur, seul le son entretenu de la voix est strictement harmonique (voir chapitre 1). La guimbarde et l'arc (pincé ou percuté) ne contredisent pas fondamentalement notre propos, car leur inharmonicité est faible dans la plupart des cas.
20. Exemple déjà présenté au chapitre 1, § 6.1. Pour nombre d'entre nous, ce type de chant est maintenant familier. Reportez-vous au chapitre 9 pour la technique vocale.

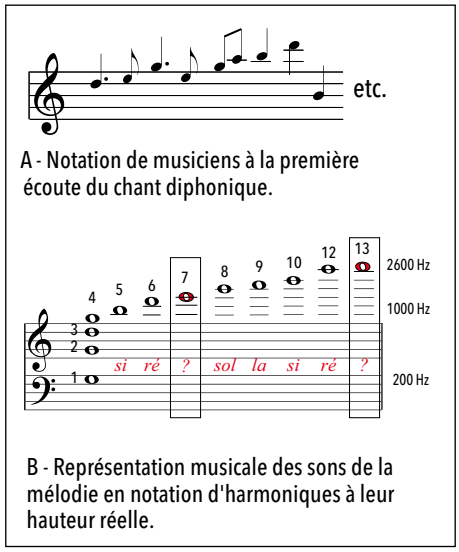


Figure 8.6 Analyse d'un chant diphonique mongol. A - notation musicale d'un musicien occidental ; B - écriture musicale des sons constitutifs de la mélodie par leur numéro dans le rang de la série harmonique et dans la tessiture réelle.

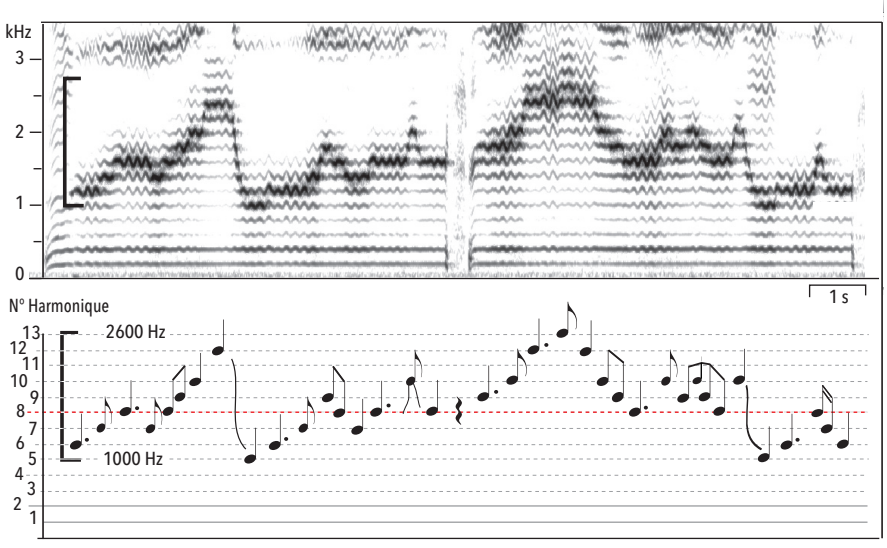


Figure 8.7 Analyse sonagraphique de la mélodie avec, au-dessous, la notation de la mélodie sur la grille des harmoniques de la voix. L'harmonique 11 n'est pas sélectionné par le chanteur.

1.3.2. Une mélodie jouée à la guimbarde

Son 8.5 (21")

Solo de guimbarde (Kazakhstan)

L'exemple suivant (Son **8.5**) enregistré par Jean During au Kazakhstan, est joué à la guimbarde (figure 8.8). Le fondamental de l'instrument est sol_1 (96,75) Hz. Le musicien commence sur H8 (sol_4) puis développe une mélodie sur H9, H10, H12 (*la, si, ré*). À deux reprises, aux temps 14,5 s et 17,5 s de notre exemple, apparaît l'harmonique 7 (fa_4) avec une sonorité étrange qui attire l'attention. En observant l'analyse spectrale (flèches sur la figure 8.9), on remarque que le son H7 est mis en évidence par le jeu de la guimbarde sans souffle, ce qui génère un spectre composé uniquement d'harmoniques impairs qui contraste avec le reste de la mélodie (voir aussi chapitre 7, le Son **7.17**).

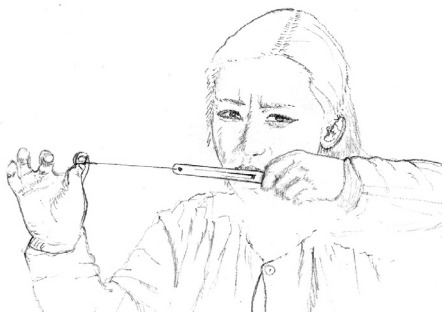


Figure 8.8 Joueuse de guimbarde (Népal)

D'après C. Jest, in Dournon-Taurelle et Wright J., 1978, p. 34.

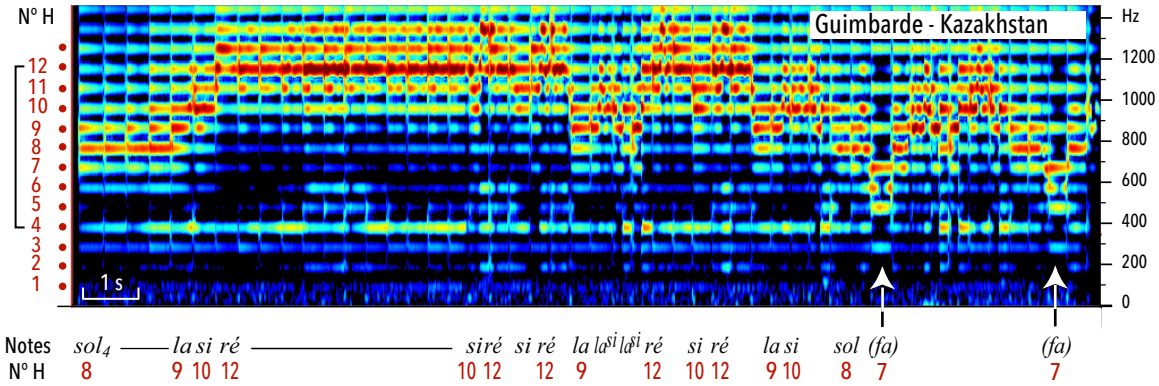


Figure 8.9 Solo de guimbarde se déroulant dans la tranche des harmoniques 8 à 12, et apparition épisodique de l'harmonique 7 en technique de jeu non soufflé.

Ces exemples illustrent une fois de plus l'importance de la perception catégorielle. Les harmoniques 6 (*ré*) et 8 (*sol*) forment bien une quarte pure, et H7 est un son « naturel », intermédiaire entre *mi* et *fa*. Un auditeur occidental attribue à ce son, incongru pour lui, l'un ou l'autre des degrés de la gamme diatonique, en fonction du contexte musical.

Un grand nombre des musiques qui se sont développées par sélection harmonique utilisent la zone comprise entre les harmoniques 6-12 (voire 6-16), car cette zone offre plusieurs intervalles intermédiaires entre les bornes fondamentales de la quarte et de la quinte. Pour préciser la grandeur de ces intervalles de façon pratique nous proposons un petit guide de calcul pages 405 et 406.

1.4. L'estimation d'un intervalle par l'harmonique commun à deux sons

Depuis qu'existe l'analyse spectrographique, la recherche d'un harmonique commun à deux sons est une méthode exploratoire simple qui permet d'estimer les intervalles avant de se livrer à la mesure des fréquences et à des calculs quelquefois fastidieux. Elle est applicable aux sons périodiques dont le spectre est suffisamment riche pour qu'il soit possible d'apprécier la coïncidence d'harmoniques de haut rang, afin d'avoir une bonne précision pour les petits intervalles dont l'harmonique commun est élevé dans la série. La méthode repose sur le principe suivant.

Lorsque l'expression numérique de l'intervalle entre deux sons est un rapport de deux nombres entiers, ceux-ci désignent le rang des harmoniques qui ont la même fréquence. Par exemple, deux sons à intervalle de quinte, dont les fréquences sont dans le rapport 3/2, sont tels que l'harmonique 3 du son grave a la même fréquence que l'harmonique 2 du son aigu : c'est l'harmonique commun.

Harmoniques communs aux deux sons d'un intervalle de la série harmonique

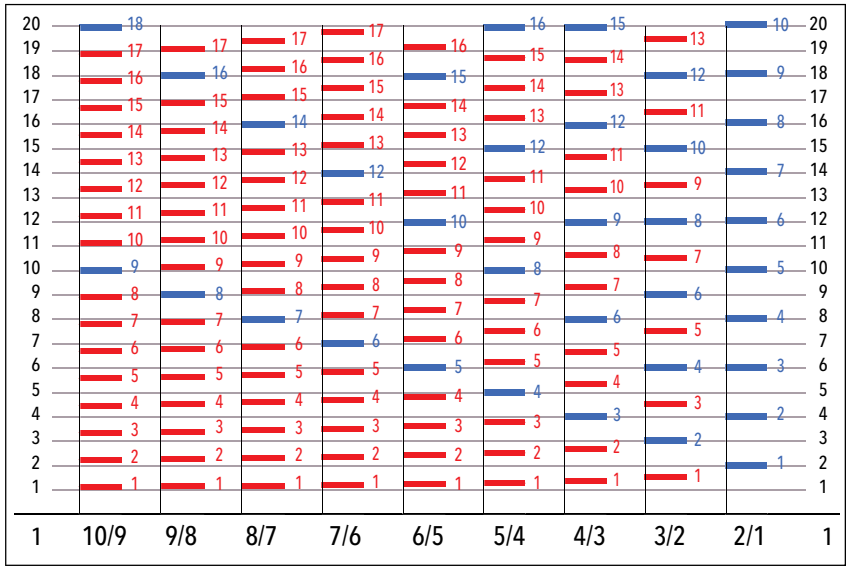


Figure 8.10 Tableau des coïncidences d'harmoniques entre deux sons. Soit un son de référence, noté « 1 », dont les harmoniques, numérotés de 1 à 20, sont figurés par des traits gris. Chaque colonne représente un son riche en harmoniques (traits rouges) dont la fréquence fondamentale est dans un rapport entier avec celle de 1 (rapport indiqué dans la partie inférieure de la figure). Les harmoniques communs entre chacun des sons et le son 1 sont indiqués en bleu. Exemple pour l'intervalle noté 10/9 : l'harmonique 9 (bleu) et l'harmonique 10 (gris) ont la même fréquence.

Le schéma de la figure 8.10 représente le spectre harmonique d'un son de référence ayant 20 composantes (grille de traits gris), et les spectres des sons qui font avec lui les intervalles des 10 premiers harmoniques successifs. Le principe est le même pour deux sons à intervalle quelconque, par exemple 8/5 ou 7/4. Lorsque le nombre d'harmoniques est grand, on a intérêt à confirmer l'estimation en observant aussi les multiples : par exemple, 14 et 12 pour une coïncidence 7/6.

Cette méthode va trouver application dans les deux exemples suivants : l'un mélodique, joué à l'arc musical, et l'autre chanté avec un intervalle en sons simultanés.

1.4.1. Exemple 1 : l'intervalle mélodique entre les deux fondamentaux d'un arc musical

Son 8.6 (23")

Jeu de l'arc musical

Figure 8.11 Joueur d'arc ngbaka : Pygmée de Centrafrique. Notez que la corde est excitée en son milieu.

D'après Castellengo, 2009, Paris.

L'arc musical présente une très grande diversité de facture et de modes de jeu. Dans celui-ci, appelé « arc-en-bouche », le musicien sélectionne les composantes du spectre en modifiant sa cavité buccale comme dans le jeu de la guimbarde (figure 8.11). De plus, il peut changer rapidement de fondamentale en cours de jeu en raccourcissant la longueur vibrante de la corde au moyen d'une baguette[21]. Il dispose ainsi de deux séries quasi harmoniques qu'il alterne à son gré, et au sein desquelles il choisit les sons qui lui conviennent (Son **8.6**).

Sur la figure 8.12, on peut remarquer le trait continu situé dans la partie supérieure du sonagramme : c'est l'harmonique commun aux deux fondamentaux de cet arc. Le trait a une certaine épaisseur car le point de raccourcissement de la corde peut légèrement fluctuer en cours de jeu, mais l'observation sur toute la longueur de la pièce confirme que l'intervalle visé par le musicien est le ton 8/7.

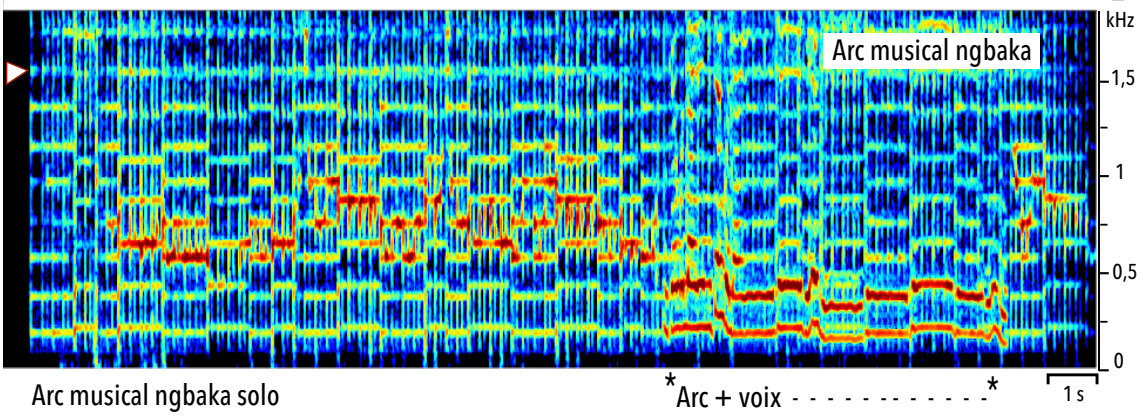


Figure 8.12 Analyse spectrographique du jeu de l'arc musical ngbaka. Le musicien alterne une mélodie jouée à l'arc et une phrase chantée, sans pour autant interrompre le jeu de l'instrument. Les deux notes fondamentales de l'arc forment un intervalle particulier, ici un ton 8/7, repérable par l'harmonique commun indiqué à la partie supérieure de la figure par un triangle blanc. FFT : 3 000 pts.

La partition de la figure 8.13 correspond à la première phrase de la mélodie, avant l'entrée de la voix. Elle représente les sons empruntés à chacun des deux sons fondamentaux de la corde qui alternent selon un rythme multiple des allers-retours de la baguette (en général 2 et 4). On voit que le joueur utilise les partiels 3-4-5 du son de la corde à vide (rouge) et les partiels 2-3-4 du son plus aigu (bleu). L'analyse détaillée des intervalles est donnée au § 1.4.4. La voix chantée entre à l'unisson de la fréquence fondamentale de l'arc (la_2 puis sol_1). Sa mélodie se situe donc deux octaves plus bas que celle qui est jouée à l'arc.

21. Voir le film réalisé par Simha Arom, expliquant la fabrication et le jeu de l'arc. Sur le site Internet : vidéotheque.cnrs.fr/doc=457.

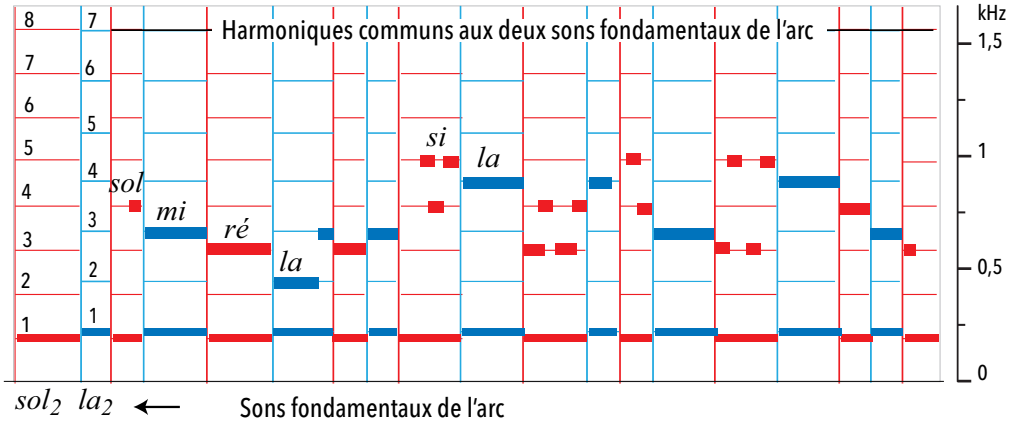


Figure 8.13 Partition schématique du début du solo d'arc de la figure 8.12. Corde à vide (sol_2 ; spectre rouge). Corde raccourcie (la_2 ; spectre bleu). L'harmonique 8 du sol_2 coïncide avec l'harmonique 7 du la_2.

1.4.2. Exemple 2 : l'intervalle harmonique entre deux voix

Dans l'exemple du Son **8.7,** deux fillettes de Côte d'Ivoire chantent en entremêlant leurs voix de façon complexe et se stabilisent à plusieurs reprises sur un intervalle sans vibrato, pendant 1 s. À chaque fois, l'intervalle ainsi chanté est une tierce. Tierce majeure ? Tierce mineure ? Autre ? À la première écoute, il nous semble entendre les deux tierces d'un accord parfait majeur, mais malgré tout l'intonation surprend. Les auditeurs ayant l'oreille absolue auront peut-être entendu *fa#-la* et *ré-fa#*, à un diapason un peu bas.

Son 8.7 (29")

Duo chanté par deux fillettes et réponses d'un chœur de femmes

Nous nous proposons de mesurer ces deux tierces, qui reviennent à six reprises dans le chant complet, toujours avec la même stabilité. L'extrait de la figure 8.14 correspond au Son **8.8.**

Son 8.8 (6")

Extrait correspondant à la figure 8.14

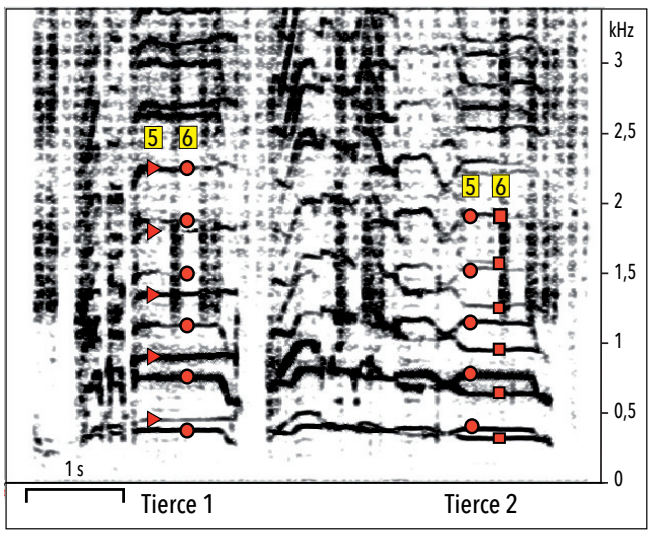


Figure 8.14 Analyse sonagraphique mettant en évidence l'harmonique commun, identique pour les deux tierces : H5 du son aigu et H6 du son grave. L'intervalle chanté est donc la tierce harmonique 6/5, intermédiaire entre les deux tierces, majeure et mineure du tempérament égal.

Les deux voix sont riches en harmoniques et se détachent bien sur le fond bruité des raclements rythmés. Bien que les harmoniques des deux fillettes soient difficiles à démêler, on remarque assez nettement la fréquence à laquelle ils se rejoignent. Cet harmonique commun se situe aux environs de 2250 Hz pour la première tierce et 1900 Hz pour la deuxième. Nous avons indiqué très précisément les harmoniques de chaque voix avec des étiquettes différentes. Une première surprise nous attend : pour chaque tierce, le 6e harmonique du son le plus grave coïncide avec le 5e harmonique du son le plus aigu. Les tierces 1 et 2 sont donc des intervalles identiques, de rapport 6/5, ce qui n'est pas aisé à entendre d'emblée.

Son 8.9 (10")

Les « notes » des deux tierces et l'intervalle de triton

L'intervalle 6/5 est une tierce mineure un peu plus grande que celle du tempérament égal. Ceci, ajouté au fait que les chanteuses alternent l'intervalle de part et d'autre d'une même note médiane, explique que plusieurs auditeurs assimilent ces intervalles aux deux tierces, majeure et mineure, d'un accord parfait. Car, lorsqu'un intervalle est entre deux catégories connues, il est facile de le rattacher soit à l'une, soit à l'autre. D'ailleurs, le calcul de l'intervalle séparant les deux sons extrêmes (note supérieure de la tierce 1 et note inférieure de la tierce 2) donne bien un intervalle de triton (608 cents) et non une quinte (700 cents)[22]. Il faut écouter cet intervalle à la fin du Son **8.9** pour se rendre à l'évidence : l'interprétation à l'oreille s'affranchit difficilement des références culturelles personnelles. Le problème est encore plus crucial pour les musiciens qui entendent des « notes ».

Cet exemple, de même que celui de l'arc musical, donne à réfléchir. L'assurance des fillettes (cinq et sept ans) est impressionnante : leur repère est manifestement celui de la coïncidence des voix sur l'harmonique commun. D'ailleurs la sonorité d'ensemble change au moment de la fusion et de la stabilisation des deux voix.

1.4.3. Réflexions sur la mesure des intervalles et sur les unités employées

On peut aussi accéder à la mesure de l'intervalle en passant par la fréquence des sons. Pour y parvenir, il faut isoler chacune des voix[23]. Nous avons mesuré, dans la partie stable, les fréquences de chaque voix. On trouve 374,6 et 449,5 Hz pour la tierce 1 ; 383 et 315 Hz pour la tierce 2. Le calcul des deux intervalles donne respectivement 315 et 339 c.

Tableau 1. Expression numérique des tierces

Intervalle	Rapport numérique	Mesure en cents
Tierce majeure du tempérament égal	1,259	400
Tierce majeure harmonique 5/4	1,25	386,28
Tierce mineure harmonique 6/5	1,2	315,61
Tierce mineure du tempérament égal	1,189	300

Comparons les valeurs des tierces des deux fillettes à celles qui figurent dans le tableau 1. On y voit : la liste par ordre décroissant des tierces majeure et mineure du tempérament égal et de la série harmonique (colonne 1) ; le quotient du rapport numérique de l'intervalle (colonne 2) ; la mesure de l'intervalle en cents (colonne 3). Les valeurs en cents des intervalles harmoniques ne sont pas des nombres entiers. Or les mesures effectuées sur des sons réels comportent nécessairement une

22. Le calcul du triton est : (6/5 x 6/5) = 36/25 = 1,419. Soit, en *cents* : $\log_2$ 2 (1,419) multiplié par 1200 = 608 c.
23. L'opération a été faite avec Audiosculpt en filtrant sélectivement toutes les composantes d'une voix sur le sonagramme, puis en effectuant la mesure des fréquences avec Praat. Ce traitement a permis de produire l'exemple du Son 8.9.

Évaluation et mesure des intervalles

Pour évaluer de façon quantitative la grandeur d'un intervalle, il faut tout d'abord effectuer une mesure. Nous prendrons comme exemple la gamme diatonique dite « naturelle » ou de Zarlin, fondée sur les rapports numériques entre les premiers nombres entiers 1, 2, 3, « consonants », auxquels a été adjoint le 5. L'explication peut aussi faire appel à la série harmonique. Les sept degrés de cette gamme sont empruntés aux sons 3, 4, 5 de la série harmonique et à leurs multiples : 4, 8, 6, 9, 15 (tableau 2).

L'intervalle de quinte *do-sol,* de rapport 3/2[1], se trouve entre les harmoniques 2 et 3 ; 4 et 6 ; 8 et 12. Le même rapport, donc le même intervalle, se trouve entre les harmoniques 6 et 9 (*sol-ré*) et les harmoniques 10 et 15 (*mi-si*). La gamme diatonique étant inscrite dans l'octave, on constate enfin que la quarte 4/3 est le complément de la quinte, et que la sixte majeure 5/3 est le complément à l'octave de la tierce mineure 6/5, ce qui permet d'introduire les degrés *fa* et *la*, étrangers à la série harmonique du *do*.

Helmholtz[2] suit une méthode plus simple basée sur le concept de tonalité. Après avoir établi que l'accord parfait majeur *do, mi, sol,* de la tonalité de *do* correspond aux sons 4, 5, 6, il fait appel aux deux tonalités relatives, à la quinte supérieure *sol majeur,* et à la quinte inférieure *fa majeur,* dont les accords parfaits donnent respectivement les sons *sol, si, ré* et *fa, la, do*. D'autres stratégies existent, mais, dans tous les cas, on aboutit à la répartition suivante des intervalles entre chacun des degrés de la gamme et le *do* de départ (tableau 3).

Notre objectif est de calculer les rapports entre sons successifs, par exemple l'intervalle entre *ré* et *mi*. Nous avons déjà *mi-do* = 5/4 et *ré-do* = 9/8. On trouve le rapport *mi-ré* grâce à un petit artifice de calcul qui consiste à introduire un troisième terme commun aux deux autres. On sait que a/b peut aussi s'écrire comme le produit (a/c) × (c/b) puisqu'en simplifiant par (c) on retrouve le rapport de départ.
On a donc : *mi/ré* = (*mi/do*) × (*do/ré*)
soit (5/4) × (8/9) = 40/36 = 10/9.

Il faut seulement prendre garde à bien écrire le rapport en respectant le sens ascendant ou descendant de l'intervalle considéré. On obtient ainsi les intervalles du tableau 4, où l'on constate que le ton a deux valeurs différentes, 9/8 et 10/9, qui sont les intervalles entre les harmoniques 8-9-10 de la tierce majeure harmonique (car 10/8 = 5/4) (tableau 4).

Pour appréhender de façon sensible la différence entre les deux tons diatoniques, par l'écoute et par le calcul, voir l'encadré page suivante.

Tableau 2. Sons de la gamme diatonique correspondant à la série harmonique

N⁰ harm.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Note	*do*	*do*	*sol*	*do*	*mi*	*sol*	---	*do*	*ré*	*mi*	---	*sol*	---	---	*si*	*do*

Tableau 3. Gamme diatonique « naturelle » : intervalles de chaque degré par rapport à la tonique *do*

Note	*do*	*ré*	*mi*	*fa*	*sol*	*la*	*si*	*do*
Intervalles	1	9/8	5/4	4/3	3/2	5/3	15/8	2

Tableau 4. Gamme diatonique « naturelle » : intervalles entre les degrés successifs

Intervalles	*do-ré*	*ré-mi*	*mi-fa*	*fa-sol*	*sol-la*	*la-si*	*si-do*
Rapport	9/8	10/9	16/15	9/8	10/9	9/8	16/15

1. Il est d'usage d'exprimer un rapport en plaçant le plus grand chiffre au numérateur, ce qui correspond au sens descendant de l'intervalle musical.
2. Helmholtz, *Théorie physiologie de la musique*, p. 22 de l'édition française, 1874.

Écoute et calculs

Son 8.10 (10")

Intervalles 10/9 et 9/8 entre do_3 et $ré_3$ (clavecin)

Le Son **8.10** permet de comparer l'intervalle do_3-$ré_3$ ajusté sur le ton mineur 10/9 puis sur le ton majeur 9/8. Cette différence paraît bien faible lorsqu'on entend les deux sons successivement. Nous verrons au §2 qu'il n'en est plus de même en contexte harmonique, en raison des battements qui se produisent sur les composantes communes.

Expression numérique d'un intervalle en *savarts* et en *cents*

La première étape consiste à effectuer les divisions. On voit que le quotient 9/8 = 1,125 est supérieur à celui de 10/9 = 1,111. Mais, pour mieux faire correspondre la mesure à la sensation de hauteur musicale, nous savons (voir chapitre 2, § 3.7) qu'il faut l'exprimer avec une unité logarithmique (base 10 ou base 2), car, dans la pratique musicale, les intervalles s'additionnent : « une tierce est l'addition de deux tons ». Le principe consiste à diviser l'octave, dont le rapport caractéristique est 2, en un petit nombre de parties égales. Deux unités existent dans la littérature : le *savart* (logarithme à base 10) et le *cent* (logarithme à base 2).

Le **savart** est la 301e partie de l'octave. En voici la définition :

$\log_{10} 2 = 0{,}30103$ et $\log_{10} 2 \times 1000 = 301$ savarts (300 savarts dans la pratique)

Le **cent** est la 1200e partie de l'octave. En voici la définition :

$\log_2 2 = 1$ et $\log_2 2 \times 1200 = 1200$ cents

Calcul des tons 9/8 et 10/9 avec les deux unités logarithmiques

L'usage des calculatrices numériques[1] a considérablement simplifié la conversion des rapports en unités logarithmiques, ce qui nous permet de donner une présentation pratique de la marche à suivre.

Avec la calculatrice, effectuer le quotient 9/8 (1,125), prendre le $\log_{10}$ du résultat et le multiplier par 1000, ce qui donne 51,15 savarts (51 sav.). Répéter l'opération en repartant du quotient 9/8, prendre le $\log_2$ et multiplier par 1200, ce qui donne 203,91 cents (204 c.). De la même façon on trouve, pour le ton 10/9, les valeurs suivantes : 45,7 sav. et 182 c.

Précision et signification du calcul

Son 8.11 (6")

Écart entre les deux $ré_3$ de l'exemple précédent (10/9 puis 9/8)

En pratique, le savart équivaut à 4 cents. Mais la question importante est celle de la précision utile. Jusqu'à quel point correspond-elle à des différences significatives pour l'écoute, et que faut-il faire des décimales ?

Le Son **8.11** donne à entendre le petit intervalle qui différencie le ton 9/8 et le ton 10/9. Cet intervalle est : (9/10) × (9/8) = 81/80 soit 5,39 savarts ou 21,50 cents. Cet exemple montre l'effet amplificateur d'un calcul effectué avec de petites unités, car « 20 » semble déjà une quantité notable alors que l'intervalle entendu est tout juste appréciable.

Lorsque nous étudions les intervalles mélodiques, les décimales sont superflues[2]. Nous dirons que ce petit intervalle – un des nombreux « commas » de la théorie musicale – est de l'ordre du dixième de ton, ce qui est très faible pour un intervalle mélodique. Le lecteur intéressé pourra pousser plus avant le calcul pour l'exprimer par rapport au ton majeur, au ton mineur ou au ton du tempérament égal qui divise l'octave en 12 parties égales. Ce ton « tempéré » vaut, par définition, 300/6 = 50 savarts ou 1200/6 = 200 cents.

La référence du tempérament égal

La division de l'octave en 12 demi-tons égaux est aujourd'hui la référence unanimement adoptée par les chercheurs pour des raisons de commodité et de neutralité : tous les intervalles ont la même valeur, quelle que soit la note de départ. Par ailleurs, la normalisation des appareils de mesure (accordeurs électroniques et numériques) et des logiciels a imposé le *cent*, très bien adapté au système tempéré. Nous aurons l'occasion d'en signaler les inconvénients pour l'étude des musiques de tradition orale, étrangères au tempérament égal (voir supra § 1.4.3).

1. Les calculatrices scientifiques sont disponibles sur la majorité des ordinateurs, mais attention : il faut veiller à choisir le bon logarithme.
2. En raison de la tolérance perceptive et aussi parce que la musique mélodique est souvent jouée par des instruments de hauteur variable.

marge d'erreur ; le moyennage statistique sur plusieurs mesures de sons variables augmente l'incertitude, et l'écoute tend à orienter l'interprétation des mesures sur des intervalles connus. Tous ces éléments expliquent les difficultés rencontrées lors de l'étude des intervalles d'une musique étrangère, surtout lorsqu'il s'agit de musique vocale. Dans notre exemple l'harmonique commun visible sur les six récurrences permet de contourner ces difficultés et surtout de s'affranchir des fluctuations de diapason au cours du chant. Il nous renforce dans l'idée que c'est bien la tierce harmonique (6/5) qui est recherchée.

Son 8.6 (23")

Jeu de l'arc musical

1.4.4. Retour sur l'arc musical : intervalles entre les sons d'une mélodie produite par décalage de deux séries harmoniques séparées par un intervalle donné

Disposons côte à côte les composantes des deux notes fondamentales de l'arc (figure 8.15). La mélodie du Son **8.6** se déroule sur les harmoniques 3, 4, 5 du son grave (sol_2, série rouge), et les harmoniques 2, 3, 4 du son de la corde raccourcie (la_2, série bleue). Le changement de fondamentale produit un décalage qui permet de jouer avec les premiers sons de chaque série, tout en disposant d'une suite de sons d'intervalles plus petits[24]. L'intervalle 8/7 entre les deux sons fondamentaux a une incidence importante sur ceux de l'échelle mélodique jouée. C'est l'intervalle formé par deux sons de même rang harmonique. Exemple : $ré_4$-mi_4 qui sont respectivement H3 du sol_2 et H3 du la_2. De même que sol_4-la_4 qui sont H4 du sol_2 et H4 du la_2.

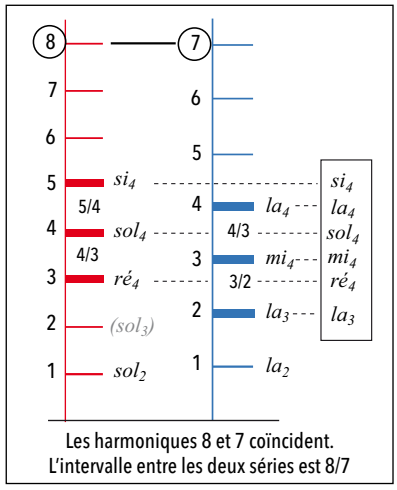


Figure 8.15 Analyse des intervalles de la mélodie jouée à l'arc musical (Son 8.6). Séries harmoniques des deux fondamentaux avec les sons effectivement joués, constitutifs de la mélodie entendue.

Les choses se compliquent lorsque le joueur enchaîne des sons de rangs différents dans les deux séries, par exemple sol_4 (H4 du sol_2) et mi_4 (H3 du la_2).

Comme nous l'avons vu au dans l'encadré page 405, on peut écrire :

$(sol/mi) = (sol/la) \times (la/mi) = (7/8) \times (4/3) = 7/6$.

L'intervalle 7/6 est intermédiaire entre la tierce 6/5 et le grand ton 8/7.

En poursuivant les opérations, nous trouvons tous les intervalles de l'échelle pentatonique de cette musique dans laquelle le son 7 joue un rôle important. On notera aussi (figure 8.16) que la quarte la_3-$ré_4$ (21/16, soit 470 cents) est un peu réduite, ce qui entraîne une quinte $ré_4$-la_4 (32/21, soit 729 cents) agrandie d'autant[25]. Ces résultats fournissent seulement un cadre intervallique de référence, car la réalisation musicale sur un instrument s'écarte toujours, peu ou prou, du modèle théorique que nous venons de développer.

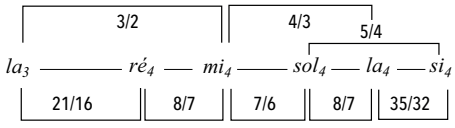


Figure 8.16 Calcul des intervalles entre les sons d'une mélodie formée par les sons de deux séries harmoniques décalées d'un intervalle de 8/7 (théorie).

24. Voir aussi chapitre 1, § 6.1, l'exemple de la tilinca roumaine.
25. Voir annexe B, § 1, le tableau des principaux intervalles exprimés en *cents* et en *savarts*.

1.5. Entre mesure et perception : les intervalles mélodiques et la musique

Comme les exemples musicaux précédents l'ont montré, les données de la perception sonore et les méthodes d'étude qui conjuguent l'analyse spectrale et la mesure des fréquences, jointes à une écoute assistée de diverses manipulations du son (ralentissement, filtrages, transposition) apportent des moyens nouveaux à l'étude des échelles musicales. Cependant, une difficulté majeure subsiste, celle de l'interprétation des mesures acoustiques au regard de la variabilité intrinsèque des réalisations musicales. L'estimation de la marge de variation acceptable pour un système donné relève, pour une part, de la tolérance perceptive des auditeurs qui n'entendent pas les petits écarts, et pour une autre, du plaisir que ces écarts procurent aux connaisseurs. Le problème est crucial en musique mélodique.

1.5.1. La variabilité des intervalles dans le jeu musical

La quasi-totalité des instruments mélodiques à vent et à cordes, et plus particulièrement la voix humaine, produisent des sons fluctuants, dont la hauteur est sous le contrôle du musicien. Les premiers chercheurs qui ont obtenu une analyse temporelle objective de la fréquence fondamentale d'une mélodie chantée ou jouée au violon ont été extrêmement surpris par l'importance des variations de fréquence affectant des notes qui paraissaient précises à l'écoute (Metfessel, 1928 ; Seashore, 1938)[26].

L'interprète d'une mélodie combine simultanément des variations de sonorité et d'intensité, et multiplie les ornements de la ligne mélodique. Outre les glissés, vibratos et diminutions diverses, les intervalles sont eux-mêmes modifiés à des fins expressives. Dans toutes les cultures, la belle exécution d'une mélodie tient moins à la réalisation précise d'intervalles codifiés qu'aux mille et une façons de les détourner, de les suggérer sans les atteindre, d'y échapper et d'y revenir. Un tel jeu musical atteint son paroxysme dans l'alliance entre un instrument de hauteur fixe (par exemple la tampoura de la musique de l'Inde) et un autre, libre de toute contrainte, comme la voix ou le violon. Sur fond d'une grille sonore immuable, les écarts du mélodiste sont perçus de façon aiguë, ils peuvent alors être très raffinés et prennent toute leur saveur.

Pour donner une idée des variations propres au jeu mélodique, nous proposons l'analyse de deux exemples empruntés à une musique dont la référence d'échelle est connue, et pour laquelle les instruments sont ajustés en conséquence.

1.5.2. Mélodie jouée à la flûte traversière : expérience en laboratoire

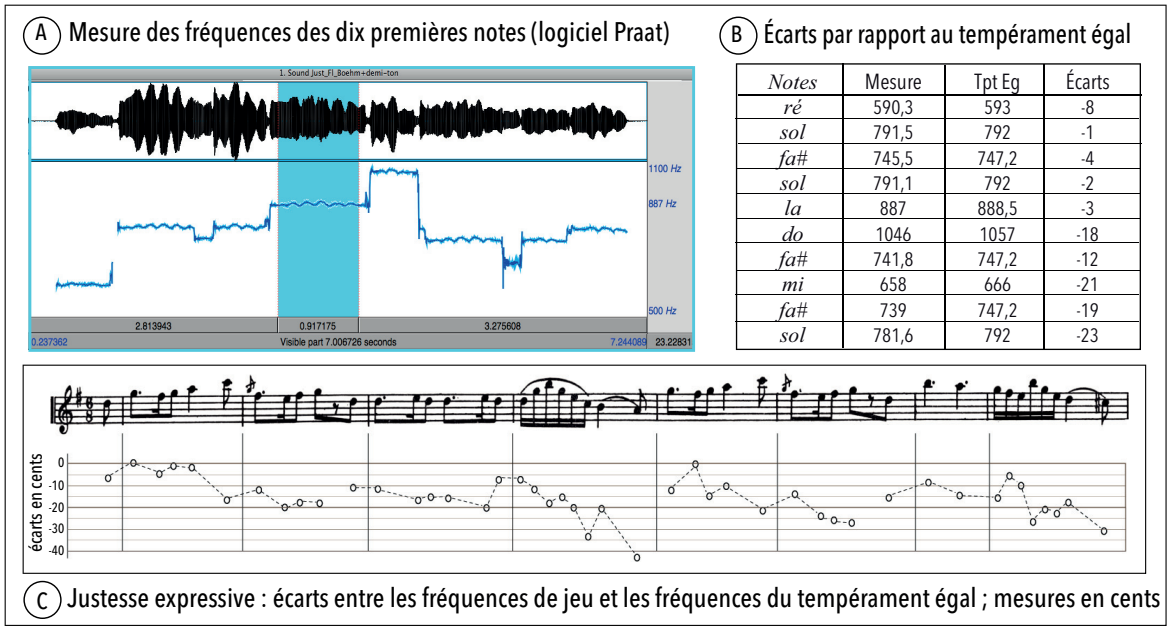
Son 8.12 (30")

Flûte traversière ; la phrase musicale mesurée figure 8.17

Partant d'une mélodie tonale extraite du *Londoner trio* n° 1 de Haydn, nous souhaitons mesurer les écarts entre l'échelle théorique, qui est la gamme à tons et demi-tons égaux, et sa réalisation instrumentale. La mélodie est jouée par un flûtiste professionnel informé des objectifs de l'expérience (Son **8.12**). Nous lui demandons de jouer avec l'expressivité musicale qu'il donne habituellement. Pour faciliter la mesure nous choisissons un mouvement modéré – *andante* – et nous le prions de réduire le vibrato. Les analyses effectuées avec Praat fournissent la courbe de fréquence moyenne affichée selon une échelle logarithmique (figure 8.17-A). Il faut ensuite calculer pour chaque degré l'écart entre la fréquence mesurée et celle de l'échelle de référence correspondant au diapason de l'instrument (figure 8.17-B). En portant les points mesurés sur un graphique en regard de la partition (figure 8.17-C), on peut interpréter les variations qui sont dues soit à des défauts de

26. Le premier tracé graphique des hauteurs effectivement réalisées par des musiciens est dû à Milton Metfessel, qui a imaginé un système automatique combinant une caméra et un oscilloscope. Voir Seashore, 1938, chapitre 26, *Primitive music, Negro songs* et p. 357 le paragraphe « Intervals ».

l'instrument (ici le do_4 est trop bas), soit à des intentions expressives plus ou moins conscientes de l'interprète : parallélisme avec le sens de la mélodie, resserrement des demi-tons, etc.



Notes	Mesure	Tpt Eg	Écarts
ré	590,3	593	-8
sol	791,5	792	-1
fa#	745,5	747,2	-4
sol	791,1	792	-2
la	887	888,5	-3
do	1046	1057	-18
fa#	741,8	747,2	-12
mi	658	666	-21
fa#	739	747,2	-19
sol	781,6	792	-23

Figure 8.17 Mesure de la justesse de jeu : flûte traversière Boehm. La référence théorique de justesse est le tempérament égal. Après que la fréquence de chaque note jouée a été mesurée, (A), les écarts sont calculés en cents (tableau B) puis tracés sur un graphique en regard de la partition (C). Les écarts observés approchent le 1/8 de ton (25 cents).

1.5.3. Mélodie jouée au hautbois : enregistrement du commerce

La figure 8.18, page suivante, montre les courbes des mesures effectuées sur un solo de hautbois[27] extrait de trois enregistrements commerciaux de *La Damnation de Faust* de Berlioz qui rendent compte d'une interprétation en contexte ordinaire de jeu[28]. Les écarts observables entre l'intervalle nominal et sa réalisation musicale sont du même ordre de grandeur que pour l'exemple de la flûte traversière : ils sont compris entre 20 et 30 cents. De façon similaire aussi, on observe pour les trois interprètes un parallélisme d'ensemble entre l'évolution de la courbe mélodique et celles des écarts par rapport à l'axe 0. Il est intéressant de comparer ces résultats avec les mesures du même extrait effectuées à l'opéra de Paris lors d'une mission sur le diapason (Leipp & Castellengo, 1965, figure 9).

Les considérations énoncées ci-dessus permettent de comprendre la résistance à l'introduction des micro-intervalles dans la composition musicale. Indépendamment des difficultés de jeu ou des problèmes de facture que ces recherches soulèvent, elles se heurtent surtout aux difficultés d'écoute d'intervalles ambigus ou trop différents de ceux de notre échelle habituelle. L'agrandissement ou la diminution des intervalles est une des ressources expressives de l'interprète pour moduler la

27. *Air du roi de Thulé*, extrait de *La Damnation de Faust* de Berlioz ; acte III, scène XIII.
28. Sources, CD 1 : London Symphony Orchestra, Dir. Colin Davis. CD 2 : Orchestre de l'Opéra de Lyon, Dir. Kent Nagano. CD 3 : Boston Symphony Orchestra, Dir. Charles Munch.

tension de son discours, et il va de soi que ces variations sont d'autant plus appréciables et appréciées que le système est connu et bien codifié. Ces questions d'écoute et de formalisation des échelles sont au cœur de l'ethnomusicologie.

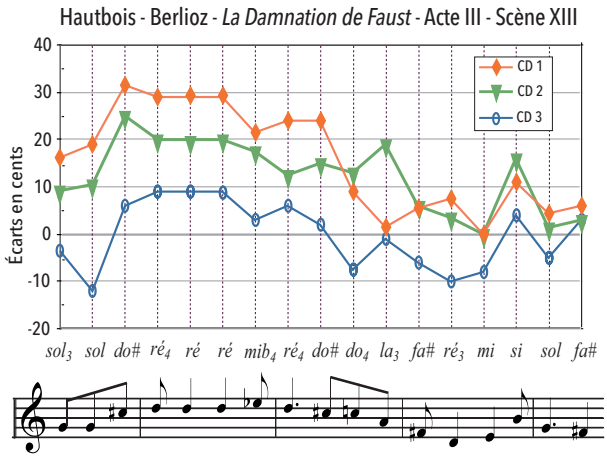


Figure 8.18 Mesure de la justesse de jeu mélodique par référence à celle du tempérament égal (axe 0 : diapason la_3 = 440 Hz). Le *do#*, sensible de dominante, est joué plus haut que la moyenne dans les trois interprétations. Remarquez la similitude entre le sens ascendant-descendant de la mélodie sur la portée musicale et l'allure générale des courbes. *Air du roi de Thulé*.

1.6. En conclusion

Les moyens d'investigation sonore et de mesure qu'offre l'acoustique ont commencé à remettre en question les conceptions ancrées dans la culture occidentale, qui proposent des systèmes de génération d'intervalles à vocation universelle. Mais la mesure n'est qu'un outil qu'il faut appliquer à bon escient, et savoir interpréter. Nous avons conclu, à la fin du chapitre 6, que la correspondance entre fréquence mesurée et sensation de hauteur ne s'appliquait valablement qu'aux sons périodiques. Les instruments de percussion métalliques (gongs, métallophones) ou en bois (xylophones) et les verges encastrées (sanzas)[29] très répandus dans nombre de musiques ont donc été laissés de côté, car ils posent des problèmes spécifiques aux acousticiens[30]. Il faudrait définir, pour chaque catégorie instrumentale, des corrections de fréquence en fonction de la tessiture et de l'intensité relative des composantes. L'exception faite pour les cordes pincées et frappées est acceptable dans la mesure où l'inharmonicité des composantes spectrales est régulière et de faible valeur, ce qui est généralement le cas des instruments de l'orchestre. Quant à la voix humaine, elle tient une place à part, tant par l'importance qu'elle a du point de vue perceptif que par sa capacité à jouer sur le double registre de la hauteur spectrale et de la hauteur tonale (voir chapitres 6 et 9). Par ailleurs la voix humaine, plus encore que les instruments à vent, est de hauteur variable.

L'interprétation des mesures est bien la partie la plus difficile de l'étude acoustique des systèmes d'intervalles, d'autant que les ressorts esthétiques de la musique monodique sont la variation, l'ornementation, les écarts qui donnent vie à la

29. La guimbarde tient une place particulière : lorsqu'elle est alimentée par le souffle, c'est un son entretenu.
30. Nous avons étudié en détail les ambiguïtés de hauteur que provoquent les sons inharmoniques au chapitre 6, § 3.

mélodie. Aisée à interpréter dans le cadre d'une structure musicale connue, la variabilité devient un casse-tête lorsque le chercheur est confronté à une culture étrangère où, le plus souvent, les notions de hauteur et d'intervalle aisément paramétrées en acoustique ne font pas sens aux oreilles de musiciens plutôt sensibles à la combinaison des formes spectrotemporelles globales, aux timbres instrumentaux et vocaux.

Enfin, l'écoute mentale et le système catégoriel de référence de la plupart des chercheurs sont solidement ancrés aujourd'hui dans le cadre de l'échelle chromatique du tempérament égal, largement étudié et commenté, doté d'un outil puissant : l'écriture sur portée. Cet outil fondamental peut se retourner contre son utilisateur, car le fait d'attribuer une étiquette à un son, même avec les corrections d'usage, lui donne une existence dans l'audition intérieure et peut conduire à des assimilations trompeuses. Il importe de toujours s'appuyer sur des documents objectifs, de lier la mesure à l'écoute en utilisant les outils de l'analyse-synthèse qui permettent d'isoler une voix, de pratiquer des transpositions, d'effectuer des ralentissements temporels afin de progressivement éduquer son oreille.

Tout ce que nous venons d'étudier sous l'étiquette générale de « mélodie » concerne les sons successifs. Il est capital de les distinguer des sons simultanés, parce que les repères objectifs d'ajustement ne sont pas les mêmes. L'évaluation d'un intervalle entre deux sons successifs s'opère dans la mémoire et se révèle très sensible au contexte, ce qui explique les tolérances observées.

Au contraire, deux sons émis simultanément produisent des phénomènes physiques objectifs, aisément repérables et très précis. C'est le domaine des sons stables, et principalement celui des instruments polyphoniques à clavier pour lesquels existent plusieurs systèmes d'accordage.

2. Sons stables simultanés : phénomènes physiques

2.1. Les battements entre deux sons voisins de l'unisson

2.1.1. Écoute mélodique et écoute polyphonique : expérience avec deux diapasons (sons quasi purs)

Prenons deux diapasons à fourche[31] vibrant à 440 Hz (la_3), et désaccordons l'un des deux en surchargeant l'extrémité d'une de ses branches avec une petite boulette de pâte à modeler. L'écart entre les deux diapasons entendus successivement est si faible que certains auditeurs n'y sont pas sensibles (Son **8.13**). Mais, dès que l'on fait sonner les deux diapasons simultanément, leurs vibrations se combinent en produisant une variation périodique de l'amplitude sonore : ce sont des *battements* qui attestent de la différence de fréquence entre les deux sources (voir Glossaire). Dans notre exemple, on compte 4 battements par seconde, ce qui signifie que le deuxième diapason, abaissé par l'ajout de pâte à modeler, est à 436 Hz.

Son 8.13 (10″)

Le diapason A puis le diapason B

31. Remarque : ce choix peut sembler contradictoire avec ce que nous venons de dire puisqu'il s'agit d'un son non entretenu. Cependant, le diapason est un instrument très particulier, dont le premier mode vibratoire produit une fréquence stable, intense et peu amortie, ce qui convient parfaitement à notre propos. Voir aussi chapitre 6, § 3.3.

Son 8.14 (14")

Le diapason A, le diapason B et les deux diapasons joués simultanément

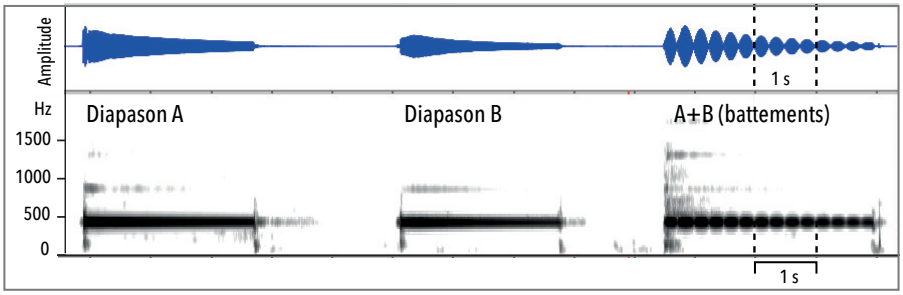


Figure 8.19 Battements de deux sons simples de fréquences voisines. Le Son 8.14 donne à entendre successivement chaque diapason frappé et posé sur une table de résonance puis les deux diapasons frappés simultanément. Les battements donnent lieu à une modulation d'amplitude visible sur la courbe supérieure et par la variation d'épaisseur de la fréquence moyenne. Leur nombre, 4 Hz, correspond à l'écart de fréquence entre les deux diapasons.

Les analyses de la figure 8.19 correspondent à la séquence sonore du Son **8.14**. En effectuant soi-même cette expérience, on pourra remarquer que les battements sont plus nettement perceptibles lorsqu'on présente les deux diapasons à une seule oreille (battements monauraux), que lorsqu'on approche un diapason de chaque oreille (battements binauraux : voir chapitre 3, page 131).

Cette expérience, facile à réaliser, fait prendre conscience du changement radical qui s'opère dans l'appréciation des hauteurs lorsqu'on passe de l'écoute mélodique à l'écoute polyphonique. L'appréciation de l'intervalle entre deux sons successifs est difficile à définir précisément, car il comporte une part d'interprétation. Au contraire, le phénomène physique des battements qui se produit lors de l'émission simultanée fournit une indication quantitative d'une grande précision. Pour réaliser un unisson parfait, il suffit d'éliminer les battements.

2.1.2. Expérience avec deux sons complexes

Son 8.15 (4")

Battements de sons complexes

Considérons maintenant deux sons périodiques riches en harmoniques. Le Son **8.15**, déjà présenté (chapitre 3, § 3.2.5), a été réalisé par synthèse en combinant deux sons de fréquences voisines : $f1 = 523$ Hz et $f2 = 527$ Hz.

On voit sur la figure 8.20, que les battements se produisent sur tous les harmoniques, et que leur rapidité est multipliée par le rang de l'harmonique. Pendant la durée d'une seconde on lit 4 battements sur le fondamental H1, $2 \times 4 = 8$ sur H2, $3 \times 4 = 12$ sur H3, etc. À l'écoute du son, il est clair que la période du battement est bien de 4/s, quelle que soit la complexité du spectre, car tous les maxima d'amplitude coïncident périodiquement avec celui de l'harmonique 1 ; voir les traits verticaux sur la figure et l'enveloppe de la courbe d'amplitude (partie supérieure).

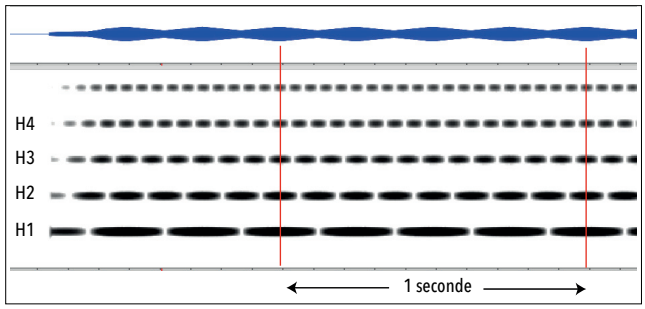


Figure 8.20 Analyse des battements produits par deux sons complexes voisins de l'unisson : courbe d'amplitude et sonagramme.

2.1.3. Perception de battements très lents avec deux sons complexes

Un curieux effet perceptif se produit lorsque l'écart de fréquence entre les deux sons est très faible, et que le battement est par conséquent trop lent pour être repérable en tant que tel.

Dans notre exemple (Son **8.16**), l'attention auditive est attirée, non plus par la synchronisation des maxima d'amplitude espacés dans le temps (ici 8,7 s), mais par les annulations d'amplitude qui affectent individuellement les harmoniques (zones blanches sur la figure 8.21) et parcourent le spectre dans un mouvement ascendant-descendant. Ce phénomène, appelé *phasing*, ne se produit qu'avec des sons extrêmement stables et riches en harmoniques comme ceux du clavecin ou des jeux d'anche de l'orgue. Il est pris en compte par l'accordeur lors de l'ajustement précis d'un unisson ou d'une octave.

Son 8.16 (15″)

Battements très lents (phasing)

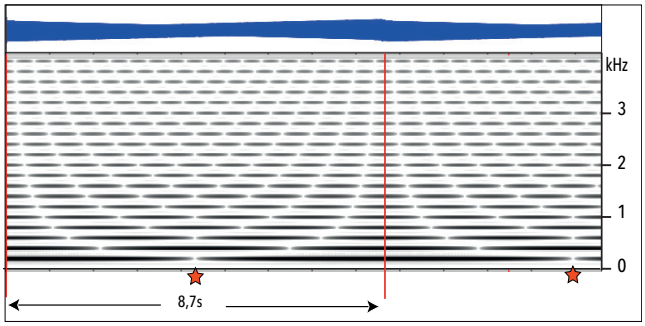


Figure 8.21 Un battement très lent (ici près de 9 secondes) produit un effet de *phasing* spectral ascendant-descendant. Remarquer l'octaviation au moment de l'annulation d'amplitude des harmoniques impairs en milieu de cycle (étoiles).

2.2. Les battements d'intervalles quelconques

2.2.1. La richesse spectrale et la perception des battements

Les battements permettent d'apprécier avec une grande précision l'écart de fréquence entre deux sons proches de l'unisson. Qu'en est-il pour de grands intervalles comme la quinte, la quarte ? L'expérience suivante, réalisée avec un instrument électronique, met en évidence le rôle des harmoniques dans la perception des battements. Il s'agit d'une mélodie chromatique en mouvement ascendant-descendant. L'exemple joué en sons purs (Son **8.17a**) ne donne pas lieu à des remarques particulières, mais l'écoute de la même séquence répétée avec des sons riches en harmoniques (Son **8.17b**) paraît soudain très fausse par endroits. L'emploi d'un instrument de synthèse permet d'affirmer que les intervalles sont exactement les mêmes en (a) et en (b), et que seul le contenu spectral a changé. Cette expérience démontre que, lorsqu'on parle de justesse, il importe de prendre en compte la composition spectrale des sons, puisque la perception des intervalles est beaucoup plus précise quand le son est riche en harmoniques.

Son 8.17a (10″)

Séquence chromatique en sons purs

Son 8.17b (10″)

Même séquence avec des sons riches en harmoniques

2.2.2. Les battements des harmoniques communs

Les intervalles les plus utilisés pour accorder les instruments de musique – octave, quinte, quarte, tierce majeure – sont ceux qui se trouvent entre les cinq premiers sons de la série harmonique. Le rapport numérique qui caractérise chacun d'eux

correspond au premier harmonique commun aux deux sons de l'intervalle (voir page 401). La figure 8.22 représente les coïncidences harmoniques de ces quatre intervalles. Si l'intervalle est plus grand ou plus petit que le rapport numérique exact, les battements vont se produire sur cet harmonique commun, ou sur ses multiples.

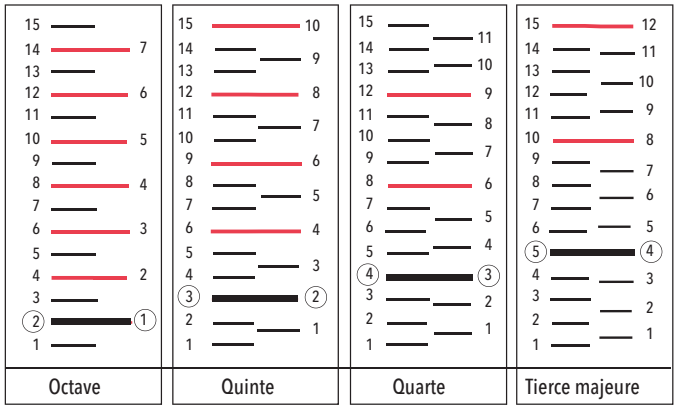


Figure 8.22 Schéma des harmoniques communs aux premiers intervalles de la série harmonique (échelle linéaire en fréquence). Le premier harmonique commun de l'intervalle est surligné par un trait noir, les autres sont en rouge.

La figure 8.23 montre l'analyse spectrale du Son **8.18**.

Son 8.18 (15")

(1) Deux sons à intervalle de quinte avec battements ; (2) même exemple avec une quinte pure (sans battements)

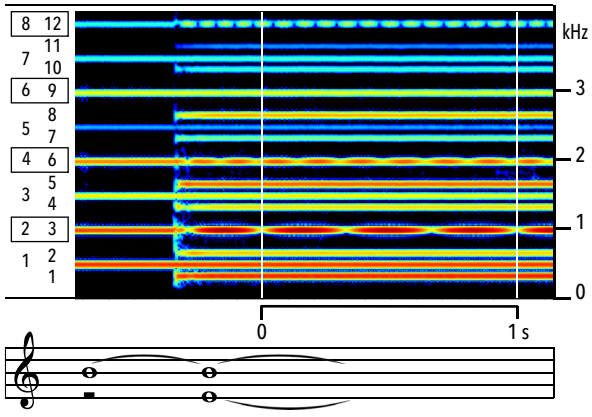


Figure 8.23 Analyse spectrale d'un intervalle de quinte joué sur un orgue électronique. Le si_3, joué en premier, est suivi du mi_3, avec lequel se produisent des battements, car ces deux sons ne sont pas dans le rapport exact de la quinte pure.

On entend tout d'abord un si_3, puis le mi_3 avec lequel se forme l'intervalle de quinte et, immédiatement, un battement de 3 Hz indiquant que la quinte n'est pas dans le rapport exact 3/2. Dans notre exemple si_3 = 486 Hz et mi_3 = 325 Hz, ce qui donne pour le premier harmonique commun : $486 \times 2 = 972$ Hz pour le si_3 et

$325 \times 3 = 975$ Hz pour le mi_3. Si nous prenons comme référence la fréquence du si_3, nous en concluons que le mi_3 est trop haut de 1 Hz, puisque l'écart de fréquence de 1 Hz sur la fréquence fondamentale du mi_3 correspond à trois battements sur le premier harmonique commun. Observez la croissance du nombre des battements avec le rang de l'harmonique commun. Dans la suite de l'exemple sonore, on entend la quinte pure sans battement avec un mi_3 accordé à 324 Hz.

Les battements jouent un rôle fondamental dans la pratique musicale : ils sont le fondement de l'accordage des instruments polyphoniques à sons fixes (clavecin, orgue, piano). Grâce à eux, le musicien dispose d'une très grande précision pour ajuster un intervalle (cordes à vide du violon et de la guitare).

2.2.3. Les effets esthétiques des battements

La recherche de battements est affaire de goût. On emploie volontairement les battements lents pour produire un vibrato d'amplitude sur plusieurs instruments. Ainsi, à l'orgue, les jeux dits ondulants comme la voix céleste (orgue français) ou la *voce umana* (orgue italien) sont obtenus par l'union de deux rangées de tuyaux accordés avec un écart de fréquence constant de l'ordre de 4 Hz (voir chapitre 3, § 3.2.2 et chapitre 6, § 4.2.1). Le même principe est exploité à l'harmonica et à l'accordéon pour l'effet appelé musette ou « brio ».

Les battements rapides, réputés produire un effet de rugosité, sont manifestement recherchés dans certaines cultures musicales. Citons la musique de Bulgarie, dans laquelle les deux lignes mélodiques simultanées de la flûte double ou des voix chantées se rapprochent et se stabilisent sur des intervalles inférieurs à la seconde mineure, qui produisent un effet saisissant de tension. Curieusement, le même goût pour les battements rapides se retrouve en Indonésie[32] et en Bolivie[33].

2.3. Des battements aux sons différentiels : « les sons ont une ombre »

Une modulation très rapide, supérieure à 50 Hz, peut donner lieu à la perception d'un son grave autonome appelé son différentiel (voir chapitre 3, § 3.2.5). Les phénomènes non linéaires qui se produisent lorsque deux sons complexes sont émis simultanément sont extrêmement difficiles à démêler. Du point de vue physique, toutes les composantes en présence se combinent entre elles, par addition et par soustraction de leurs fréquences, en produisant les sons dits résultants[34]. Et, comme l'écrit Gérard Grisey : « Ceci nous amène à considérer qu'autour de chaque intervalle et de chaque complexe de sons, se trouve une aura de sons différentiels et additionnels dont la complexité dépend de la richesse en harmoniques des sons générateurs : *les sons ont une ombre.* »[35]

Dans la réalité seuls les sons différentiels sont clairement audibles, et seulement dans certaines conditions : lorsqu'on combine deux sons de forte intensité, et de préférence dans une tessiture élevée.

32. Rappoport, D., 2011, *Chant de la terre aux trois sangs* ; livre + DVD, éditeurs MSH-Épistèmes, Paris.
33. Citons la flûte *taraka*, Gérard-Ardenois, D., 2002.
34. Voir Helmholtz, 1874, chapitre VII, p. 192-198.
35. Grisey, G., *Structuration des timbres dans la musique instrumentale*, 1991, collectif-Barrière, p. 369.

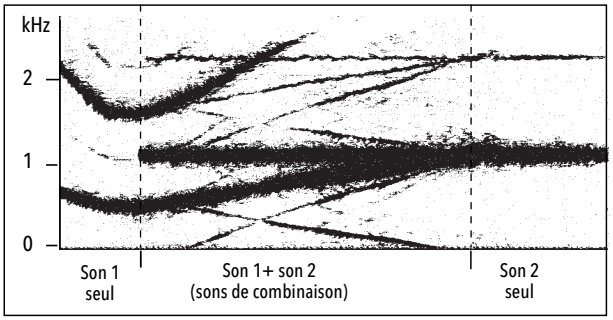


Figure 8.24 Sons de combinaison, additionnels et différentiels, tels qu'ils apparaissaient à l'analyse avec une machine analogique (Sona-Graph analyzer) antérieure à 1990. Ces phénomènes ne sont plus visibles sur les analyses numériques.

Pour en faire l'expérience, on peut prendre les embouchures de deux flûtes à bec soprano dont la fréquence est environ 1 600 Hz (sol_5), les emboucher ensemble et souffler assez fort. Il faut ensuite abaisser le son de l'une d'elles en recouvrant partiellement la partie inférieure du tube avec une main, l'autre embouchure restant dégagée. On a alors la surprise d'entendre un troisième son variant en sens inverse du premier (figure 8.24).

Les sons différentiels ont été décrits dès le XVIIIe siècle par le violoniste G. Tartini[36]. Leur existence objective est toujours contestée, les chercheurs attribuant leur audibilité aux distorsions de l'oreille (Pressnitzer & Patterson, 2002). Ce phénomène est tout de même remarquable dans la tessiture mentionnée et pour des instruments comportant beaucoup d'énergie dans le fondamental. Ils peuvent être sciemment utilisés en musique comme dans l'exemple du duo de flûtes de la figure 8.25.

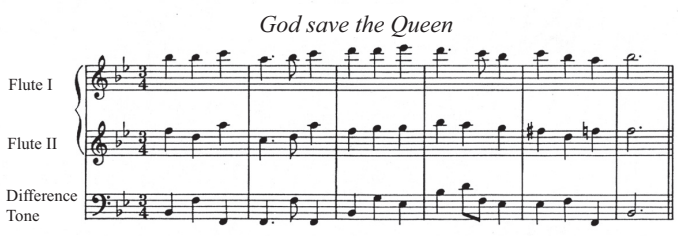


Figure 8.25 Duo pour flûte avec notation de la mélodie résultante due aux sons différentiels.

D'après Wye, T., *Practice book for the flute*, vol. 4, Intonation and vibrato, Éd. Novello.

Un tel exercice est difficile, car il exige de jouer sans vibrato et d'ajuster l'intervalle avec une grande exactitude, faute de quoi le son différentiel est plus haut ou plus bas que la note indiquée, ce qui donne un résultat plutôt désagréable à entendre. Comme son nom l'indique, la fréquence d'un son différentiel est strictement la différence en fréquence des deux sons primaires, ou générateurs.

Les sons différentiels peuvent rendre service dans l'orgue. Lorsque la place manque pour disposer de grands tuyaux dans un instrument, le facteur peut réaliser des jeux très graves en combinant deux tuyaux différents sonnant à intervalle de quinte. Ainsi, pour obtenir un fondamental de 60 Hz, on fait jouer deux tuyaux accordés respectivement à 120 et 180 Hz. Un tel jeu porte le nom de *basse acoustique*. L'intensité du fondamental n'est pas perçue aussi fortement que s'il s'agissait d'un tuyau réel de 60 Hz.

36. Dans un ouvrage publié en 1754. Voir aussi J.-J. Rousseau, *Dictionnaire de musique*, entrée *Système*, et planche G, figure 5.

2.4. La production de hauteurs complexes avec des instruments à sons entretenus

Parmi les différentes recherches visant à renouveler le jeu instrumental au XXe siècle, on peut citer la pratique du chant dans un instrument à vent, et les sons *multiphoniques*. Ces techniques, bien connues des musiciens traditionnels, étaient soigneusement évitées tant que l'écriture musicale privilégiait la production d'une « note » non ambiguë.

2.4.1. Chanter en jouant d'un instrument à vent

Lorsqu'un musicien chante en jouant d'un instrument à vent, il combine étroitement les fréquences fondamentales de deux sources – la voix et l'instrument –, car l'air alimentant l'instrument est déjà périodiquement modulé par la fréquence fondamentale de la voix. Le son produit, qui rappelle certains effets de modulation d'amplitude et de fréquence en synthèse sonore, est extrêmement complexe. Il dépend à la fois du contenu harmonique de chaque source : voix et instrument, et de l'intervalle entre les deux fondamentaux. Un musicien entraîné peut choisir de produire des battements en chantant une fondamentale très proche de celle de l'instrument, ou un intervalle précis écrit par le compositeur[37]. Pour effectuer un premier essai, utilisez une flûte à bec, comme dans l'exemple analysé sur la figure 8.26, et jouez une mélodie à la flûte en chantant une note tenue, de préférence plus grave, avant de vous entraîner à changer la hauteur du son vocal.

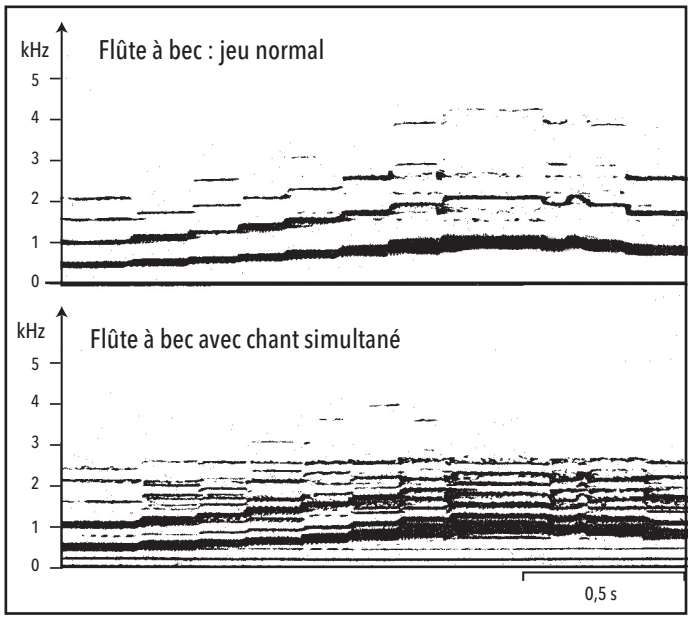


Figure 8.26 Analyse d'une gamme diatonique jouée à la flûte à bec (do_4-do_5) et de la même séquence accompagnée d'un bourdon vocal simultané sur do_3.

37. Exemples musicaux à écouter : *Maya* de Y. Taïra, pour flûte basse ; le début de *Solo pour deux* de G. Grisey, où le tromboniste chante dans son instrument. Déjà en 1806, C. M. von Weber écrit dans le *Concertino en mi mineur pour cor,* une cadence avec des accords de trois sons, qu'un corniste virtuose, Vivier, produisait en chantant dans l'instrument. Voir *Bulletin du GAM* n° 74, p. 5-7, et Wikipedia : *Cor d'harmonie*, § « Sons multiples ».

La pratique du chant couplé au jeu d'un instrument à vent, en particulier la flûte, est répandue dans le répertoire musical traditionnel de nombreuses cultures. Citons entre autres : l'Inde (Rajasthan), la Bulgarie (flûte double), le Maghreb (jeu du ney).

2.4.2. Le paradoxe des sons multiphoniques

On appelle « son multiphonique » un son émis par un instrument monodique entretenu donnant la sensation d'un accord complexe dans lequel on peut distinguer plusieurs hauteurs de sons. Par définition, la note fondamentale émise par un instrument entretenu résulte du couplage entre la fréquence d'un système oscillant (jet d'air, anche) et celle d'un mode vibratoire du tuyau. Un tel son, stable et périodique, donne la sensation d'une hauteur unique. Alors comment peut-on jouer deux ou trois notes simultanées sur un instrument comme la flûte ou le hautbois ?

Les sons multiphoniques ont été décrits en acoustique comme des phénomènes paradoxaux (Bouasse, 1929[38]), alors qu'ils font partie intrinsèquement de plusieurs musiques traditionnelles, en particulier en Bolivie et au Chili. Ils ont été introduits dans l'écriture musicale dès la deuxième moitié du XXe siècle (Bartolozzi, 1967). Leur développement est dû aux recherches des instrumentistes qui ont recensé les doigtés praticables et travaillé les techniques de souffle permettant de les émettre avec sûreté, en particulier dans le cadre d'un programme de recherche (ARI) mené à l'Ircam. Citons P.-Y. Artaud à la flûte ; V. Globokar, B. Sluchin au trombone ; D. Pateau au hautbois ; D. Kientzy au saxophone ; A. Ouzounoff au basson.

Pour produire un son multiphonique, il faut en quelque sorte détourner la technique classique de jeu pour contraindre l'instrument à entretenir non pas un, mais deux ou trois modes vibratoires différents, simultanément. Les fréquences propres de ces modes n'étant pas en rapport harmonique, plusieurs cas peuvent se présenter (Castellengo, 1982) :

- l'instrument produit un son chaotique difficile à stabiliser,
- le son est affecté d'interruptions plus ou moins régulières (roulements),
- le son se stabilise sur un accord complexe : c'est ce dernier cas qui nous intéresse ici.

L'exemple **8.19** fait entendre deux sons multiphoniques stables, joués l'un sur une flûte traversière, l'autre avec un hautbois.

Son 8.19 (5")

Les deux sons multiphoniques de la figure 8.27

L'analyse spectrographique du son multiphonique de flûte traversière est extrêmement curieuse pour qui connaît bien cet instrument. On observe sur la figure 8.27 une grande quantité de composantes apparemment harmoniques d'un son grave dont le fondamental manque, ce qui est invraisemblable, et les composantes les plus intenses sont au milieu du spectre. En repérant les fréquences de ces composantes, nous constatons qu'elles correspondent plus ou moins aux sons notés à l'oreille : un si_3 (vers 500 Hz), un mi_4 (vers 660 Hz) et un $ré_5$ très intense (vers 1100 Hz). Le son de hautbois présente un aspect similaire. L'explication est la suivante.

Lorsqu'un son multiphonique se stabilise, les partiels des modes propres du tuyau qui sont sollicités s'accommodent entre eux de manière à entrer dans la série harmonique d'un son plus grave dont nous avons tracé la grille harmonique (rouge) sur la figure 8.27, sous forme d'une échelle verticale à la droite de l'analyse du son.

38. « Des partiels non harmoniques peuvent être émis simultanément ; bien que due à une lame d'air unique, l'excitation n'est manifestement pas périodique. », *Tuyaux et résonateurs,* § 60. Autre : « Il semble naturel de poser que l'écoulement d'une lame d'air est périodique. Or les tuyaux à cheminée donnent des accords manifestement faux. », *Instruments à vent,* T. I, § 101.

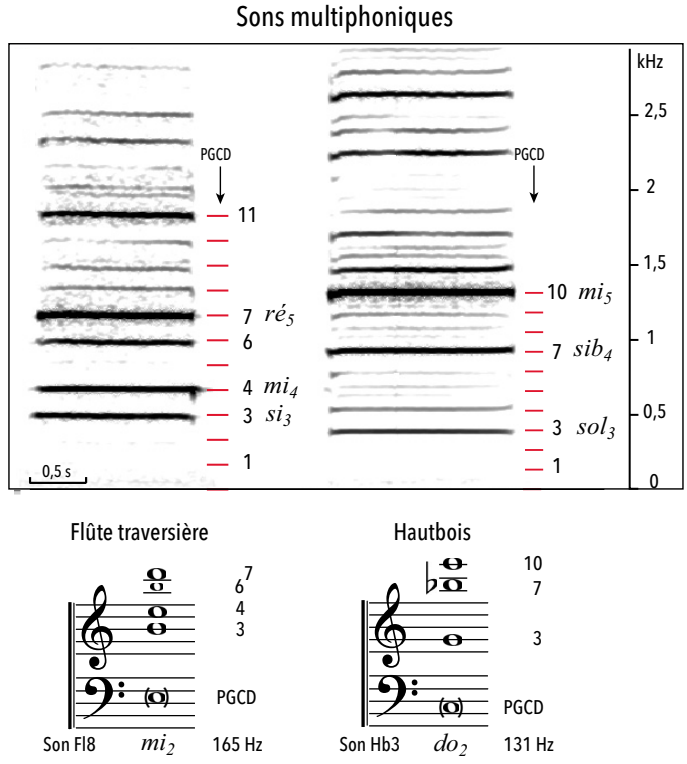


Figure 8.27 Analyse sonagraphique de deux sons multiphoniques joués à la flûte et au hautbois et notation des composantes saillantes perceptivement, considérées comme harmoniques d'un fondamental : le PGCD. Sous le sonagramme, notation des principales composantes des sons Fl8 et Hb3 perçues par les musiciens

Le fondamental de ce son grave est le PGCD[39] des fréquences émises. Mais, contrairement au modèle harmonique évoqué au chapitre 6 (voir page 255), le PGCD n'est pas perçu comme une hauteur musicale : il représente seulement la fréquence de synchronisation qui assure la quasi-périodicité du phénomène. C'est pourquoi nous avons proposé de désigner musicalement les composantes d'un son multiphonique par leur numéros dans la série harmonique du PGCD, ici 165 Hz (mi_2) pour la flûte traversière et 131 Hz (do_2) pour le hautbois. En conséquence, l'intervalle mi_4-$ré_5$ entre les sons 4 et 7 de la flûte est plus faible que la septième mineure du tempérament égal, puisqu'il s'agit de l'intervalle de 7e harmonique. Certains musiciens peuvent d'ailleurs l'entendre comme une sixte majeure, *mi-do#,* un peu grande. Ce fait, joint à la tessiture généralement aiguë des composantes dominantes de ces types de sons, explique les difficultés rencontrées par les instrumentistes et par les compositeurs pour noter exactement un son multiphonique[40].

L'exemple sonore **8.20** donne à entendre les notes principales de chaque multiphonique extraites par filtrage : composantes 3, 4, 7 pour la flûte et 3, 7, 10 pour le hautbois[41]. Cette synthèse restitue assez bien la couleur accordale de l'agrégat de sons perçus, indépen-

Son 8.20 (5")

Sons multiphoniques réduits à leurs composantes principales

39. L'abréviation PGCD signifie : plus grand commun diviseur. Le logiciel Praat en donne la mesure directe (164,7 Hz pour Fl8 ; 130,6 Hz pour Hb3).
40. Le Son Fl8 analysé correspond au son 78 de la flûte en *ut* noté dans *Flûtes au présent*, Artaud & Geay, 1980.
41. Cet exemple est extrait d'une recherche collective développée à l'Ircam. Un algorithme de description spectrale des sons multiphoniques a été développé par G. Assayag, permettant à C. Malherbe d'intégrer les sons multiphoniques dans sa pièce *Non sun* (1984). Voir Assayag & coll. (1985).

damment de la qualité sonore globale. L'effet perceptif change d'ailleurs selon les auditeurs, mais on notera, une fois de plus, le rôle important joué par la zone d'écoute dominante (500 à 1800 Hz) dans la perception de hauteur (voir Glossaire).

2.5. Les hauteurs d'un son complexe : des notes ? un accord ? un timbre ?

Depuis que nous avons accès à l'analyse des sons réels et que nous pouvons créer des sons par synthèse, la relation entre la fréquence d'un son et sa hauteur perçue, classiquement présentée en introduction des ouvrages d'acoustique, a été singulièrement ébranlée. De façon similaire, les compositeurs ont depuis longtemps contesté les frontières admises entre perception des hauteurs et timbre. L'usage grandissant de sons ambigus comme ceux des cloches, les sons multiphoniques, les accords et agrégats de sons simultanés sans relation harmonique, a jeté le trouble dans les catégories traditionnelles qui distinguent la note, l'accord et le timbre.

Son 8.21 (10")

Début de Mutations *(figure 8.28)*

Dans sa pièce *Mutations* (1969), J.- C. Risset a créé à l'ordinateur un exemple sonore intéressant, qui combine différemment les cinq sons suivants : *ré*$_4$, *do*#$_3$, *sib*$_3$, *la*$_4$, *mi*$_4$ (figure 8.28). Il nous invite avec le Son **8.21** à écouter successivement une mélodie (A), une harmonie (B) et un timbre (C). R. Erickson, qui analyse cet exemple[42], remarque que la perception de tels phénomènes change selon le contexte musical, autrement dit selon ce qui précède et ce qui suit. Il est vrai qu'un même auditeur peut, selon son bon plaisir, décider d'entendre tour à tour les notes isolées, la fonction harmonique d'un accord ou apprécier globalement le timbre d'un agrégat sans chercher à le décomposer. Cet exemple a fait l'objet de nouvelles réflexions (voir Risset, 2004, page 155 et Lévy, 2008).

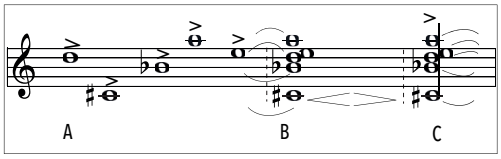


Figure 8.28 Séquence synthétisée à l'ordinateur par J.-C. Risset : cinq sons forment une mélodie, une harmonie, un timbre.

3. L'accordage des instruments polyphoniques de hauteur fixe

Les instruments polyphoniques de hauteur stable comme l'orgue, le clavecin et les instruments à frettes, sont des instruments à sources multiples indépendantes dans lesquels les intervalles, entendus en sons simultanés, peuvent être accordés précisément sur la base des battements. L'abondante littérature et les discussions qui ont lieu aujourd'hui encore sont en grande partie dues au fait que cette matière se prête à des cogitations numériques, qui entraînent certains de leurs auteurs à vouloir justifier par le calcul une solution générale censée être meilleure que toutes les autres.

Pour comprendre la complexité réelle des questions d'accordage, l'existence d'une grande quantité de systèmes et les jugements sur la justesse qui en dépendent, il faut replacer chaque système d'accordage dans son contexte musical d'origine et surtout prendre en compte les particularités acoustiques des instruments auxquels

42. Voir Erickson, *op. cit.*, chapitre 2, « Some territory between timbre and pitch », p. 19-20.

il s'applique. La présentation succincte des principes de l'accordage qui suit doit l'essentiel de son contenu au travail de Pierre-Yves Asselin (1983, 1985), organiste, chercheur et accordeur de divers instruments à clavier, dont nous avons suivi l'enseignement théorique et pratique dans la décennie 1980.

3.1. Les bases acoustiques de l'accordage

3.1.1. Un vocabulaire pour l'accordage

- **Gammes et échelles.** Les degrés d'un système musical et les intervalles qui les séparent constituent *l'échelle,* ou la *gamme,* qui le caractérise. Nous emploierons les deux termes indifféremment, sans nous rallier aux auteurs qui posent que la gamme se reproduit à l'octave, puisque les gammes dites de Pythagore et de Zarlin génèrent une infinité de sons. Plus que les noms des degrés (qui réapparaissent à chaque cycle avec des positions différentes), c'est le choix des intervalles d'une musique qui nous intéresse ici.
- **Systèmes musicaux et tempéraments.** Les systèmes d'accordage des instruments polyphoniques sont nombreux[43]. Certains sont dits *ouverts*, car ils ont un nombre limité de degrés et échappent à la contrainte de l'octave ; d'autres permettent de choisir la position d'un degré en cours d'exécution, par le biais d'un clavier particulier ou en modifiant la hauteur de jeu. Les systèmes d'accord d'instruments à sons fixes et à 12 sons dans l'octave sont nécessairement des *tempéraments,* en raison des compromis qu'il faut faire sur l'ajustement de certains degrés pour rester dans le cadre de l'octave pure.
- **Intervalles purs et intervalles justes.** On appelle *intervalles purs* ceux dont le rapport des fréquences est strictement celui de deux nombres entiers, comme ceux que l'on trouve dans la série harmonique. Par définition, ils sont donc sans battement. Un *intervalle juste* est celui qui correspond à un système de référence donné, qu'il ait ou non des battements.
- **La partition.** En technique d'accordage, la *partition* est l'opération qui consiste à répartir les 12 sons de l'échelle chromatique de sorte que le 13e soit à l'octave du premier. C'est l'opération initiale et fondamentale de l'accord dans un tempérament. Tous les auteurs recommandent d'établir la partition dans la partie médium du clavier, de sorte que les battements s'entendent bien. À l'orgue, la partition se fait sur le prestant, un jeu de 4' plus facile à entendre[44]. On étend ensuite la partition vers l'aigu et vers le grave, par octaves pures, en les ajustant par l'écoute des quintes et quartes.
- **Intervalles complémentaires.** La position d'un son dans l'octave détermine deux intervalles complémentaires. Prenons l'exemple du couple quinte-quarte : sol_3 est à la quinte supérieure de do_3 et à la quarte inférieure de do_4. Si la quinte est diminuée, la quarte, qui lui est complémentaire, est agrandie d'autant. Dans la pratique de l'accordage, on peut se baser sur l'un ou l'autre des intervalles complémentaires.

3.1.2. La pratique des battements

- **Battements et tessiture.** La *rapidité*[45] *des battements* (différence en Hz entre deux composantes voisines) sur laquelle se règle l'accordeur à l'oreille pour estimer la réduction ou l'agrandissement d'un intervalle dépend de la fréquence des sons, donc de la tessiture. Considérons par exemple une quinte légèrement réduite par

43. Plusieurs traités voient le jour dès le XVIe siècle. Voir en bibliographie Bougeret (1982), Psychoyou (2003), Lindley (1984), Asselin (1985).
44. Le *do* du milieu du clavier d'un jeu de 4' est à la hauteur du do_4, soit environ 500 Hz. Tous les sons sonnent à l'octave supérieure du son écrit.
45. Nous préférons substituer « rapidité » à fréquence des battements, pour qu'il n'y ait pas de confusion avec la fréquence des sons que l'on accorde.

rapport à la quinte pure, qui bat deux fois par seconde. À l'octave supérieure, la quinte réduite d'une quantité équivalente bat quatre fois par seconde, car toutes les fréquences sont doublées, donc l'écart de fréquence entre les harmoniques communs qui produisent les battements l'est également.

- **Perception des battements.** Si l'intervalle que l'on accorde est très proche de l'intervalle pur correspondant, il est difficile de savoir s'il s'agit de battements *par excès* (intervalle trop grand) ou de battements *par défaut* (intervalle trop petit), car la modulation d'amplitude est perçue de façon similaire. Il faut donc bouger un des deux sons de l'intervalle pour connaître la suite à donner. Ce point est fréquemment source d'erreurs.
- **Battements de sons quasi harmoniques.** Les sons quasi harmoniques comme ceux du piano, et ceux de la harpe, posent des problèmes différents selon qu'il s'agit du grave ou de l'aigu. L'accordage à l'oreille permet de tenir compte des particularités de la perception humaine.

3.1.3. Les processus d'accordage : découverte des commas

Les intervalles sur la base desquels on peut engendrer les degrés d'une échelle musicale sont la quinte et la quarte. La première étape d'ajustement d'un intervalle consiste à éliminer les battements, comme le fait un violoniste lorsqu'il accorde son instrument. Lorsque les deux sons se trouvent dans un rapport harmonique exact, une sensation particulière de calme se produit, et le timbre perçu change dès que les deux sons fusionnent. Il est possible que le premier différentiel, se trouvant aussi en rapport harmonique, participe de l'effet ressenti. Les intervalles purs sont donc aisément repérables et manifestement recherchés dans la plupart des musiques polyphoniques, y compris les musiques vocales qui réalisent dans cette esthétique une stabilité vocale assez surprenante.

Pour exposer simplement les données fondamentales de l'accordage, nous prendrons comme exemple la musique occidentale pour clavier, qui a fait l'objet d'un grand nombre de réflexions au cours de son histoire. Les problèmes acoustiques soulevés, communs à tous les systèmes musicaux, y sont clairement identifiables en raison même de l'emploi de sons fixes. On peut les découvrir en réalisant les exemples proposés sur un clavecin[46] ou sur un synthétiseur.

Lorsqu'on s'exerce à l'accordage, on découvre rapidement trois incompatibilités polyphoniques entre les intervalles purs, qui ont donné lieu à la caractérisation de commas[47]. Prenons un exemple concret réalisé sur un instrument électronique très stable et facile à accorder, le Cantor, construit au laboratoire d'acoustique[48].

Quintes pures et tierces pures : le comma syntonique (Son 8.22)

Son 8.22 (27")

Séquence d'accordage démontrant le comma syntonique ; orgue électronique *(voir page 440)*

Partant du do_3 (figure 8.29 a), nous accordons le *sol* à la quinte supérieure en éliminant les battements (1), puis le *ré* à la quarte inférieure du *sol* (2), toujours sans battement. La quinte pure supérieure du *ré* donne le *la* (3), suivi de la quarte inférieure le *mi* (4). Le résultat est agréable à écouter mélodiquement, mais pose un problème en polyphonie. En effet, les deux sons simultanés de la tierce majeure, *do-mi,* produisent des battements : ce n'est donc pas un intervalle pur. Il faut baisser légèrement le *mi* pour obtenir une tierce sans battement avec le *do*. L'écart entre le *mi* issu des quatre quintes (ou *mi* « pythagoricien ») et le *mi* tierce harmonique de *do* (en rouge sur la figure), est un petit intervalle appelé *comma syntonique* (figure 8.29 b). Une autre façon de procéder (figure 8.29 c) consiste à cumuler quatre

46. Le clavecin se prête à des essais d'accord, contrairement au piano qu'il ne faut en aucun cas se risquer à accorder, sous peine de déboires.
47. Terme grec employé traditionnellement en musique pour désigner un très petit intervalle. Il existe plusieurs sortes de commas.
48. Voir Leipp & coll., 1971, *Bulletin du GAM* n° 56.

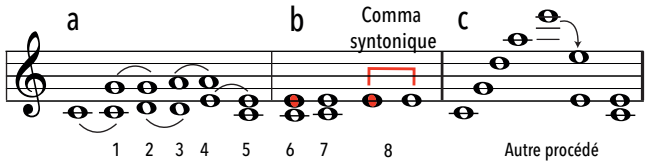


Figure 8.29 Le problème du comma syntonique. À l'issue de l'accord de quatre quintes et quartes pures sur les sons *do, sol, ré, la, mi,* la tierce majeure résultante (*do-mi*) est plus grande que la tierce pure harmonique. L'écart entre le *mi* obtenu et celui de la tierce pure (en rouge) est appelé *comma syntonique*.

quintes pures ascendantes, ce qui mène au mi_5. Il faut ensuite accorder mi_4 et mi_3 par octaves pures descendantes, ce qui conduit au même résultat.

Nous avons maintenant à notre disposition les sons *do, ré, mi, sol, la* et leurs octaves, ce qui permet de jouer un nombre considérable de mélodies pentatoniques. Cette tentative d'accord achoppe donc sur une incompatibilité fondamentale entre quintes pures et tierces pures, car si nous abaissons le *mi* pour obtenir une tierce majeure pure, alors la quinte (trop courte) et la quarte (trop grande) ont des battements.

Ce petit comma, qui peut paraître bien faible à certaines oreilles, est le « grain de sable » de l'accordage.

Pour rendre sensible le phénomène, P.-Y. Asselin a réalisé au clavecin un exemple éclairant (qui peut toutefois malmener les oreilles sensibles). Soit la séquence de cinq accords (figure 8.30) que nous voulons jouer sur un clavecin en intonation pure intégrale : quintes et tierces sans battement. Si l'on conserve la hauteur des notes communes à deux accords successifs, l'écoute (Son **8.23**) produit un curieux malaise dû au fait que le 5e accord, identique au 1er, se retrouve plus bas (d'un comma syntonique). Poursuivons l'expérience en parcourant un nouveau cycle à partir de l'accord final, le décalage s'accumule pour aboutir, au bout de 5 cycles, un demi-ton plus bas que le point de départ ! (Son **8.24**). Que se passe-t-il ? Le problème est dû au fait qu'à un moment donné de l'enchaînement des accords, un son en position de tierce, conservé comme note commune, se retrouve en position de quinte dans l'accord suivant. L'explication s'appuie sur le diagramme de la figure 8.31 à droite (voir aussi Asselin, 1985, page 130).

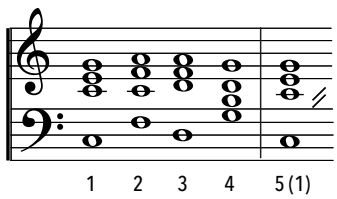


Figure 8.30 Séquence des cinq accords

Son 8.23 (8")

La séquence de la figure 8.30, jouée au clavecin (intonation pure)

Son 8.24 (41")

Effet de cinq répétitions de l'exemple précédent *(voir page 441)*

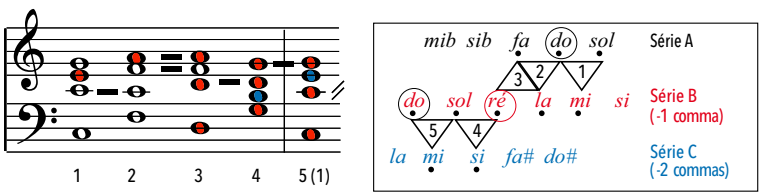


Figure 8.31 À gauche : mise en évidence des notes communes à deux accords successifs. À droite, diagramme montrant trois séries de sons (A, B, C) comportant une succession de quintes pures décalées l'une par rapport à l'autre d'un comma syntonique Les notes des accords y sont positionnées au cours des enchaînements successifs. Les couleurs des notes sur la portée musicale de gauche correspondent à celles des séries à droite.

Chaque accord parfait est figuré par un triangle dont la base est l'intervalle de quinte et la pointe le son de la tierce. Ce dernier est nécessairement emprunté à une série de sons différente. Le premier accord prend la quinte *do-sol* en série A ; le *mi*, tierce pure du do, relève d'une série de sons situés un comma syntonique plus bas (série B, rouge). Il en est de même pour l'accord 2 (*fa-la-do*). L'accord 3 s'enchaîne en conservant les sons *fa* et *la*. Le *la*, qui se trouve dans la série rouge, entraîne donc l'abaissement du *ré* qui en est la quinte inférieure : c'est le point critique. Les enchaînements suivants par notes communes – *ré* pour les accords 3 et 4, puis *sol* pour les accords 4 et 5 – tous deux situés dans la série B, conduisent alors inéluctablement à l'accord final dont la fondamentale, *do*, se trouve un comma plus bas que le *do* de départ. Le décalage se reproduit à chaque tour du cycle. Le clavecin étant un instrument à sons fixes, l'exemple a été réalisé par montage, car il a fallu réaccorder les notes à chaque répétition.

Le même problème se pose en musique vocale, cependant il passe souvent inaperçu parce que les chanteurs entraînés rajustent instinctivement les sons pendant les transitions critiques, faute de quoi l'ensemble dériverait de façon similaire en perdant le diapason de départ. Écoutez le Son **8.25** ; la figure 8.32 donne le détail de la réalisation.

Son 8.25 (21")

Quatuor vocal interprétant l'enchaînement de la figure 8.32 avec rajustement instinctif du ré

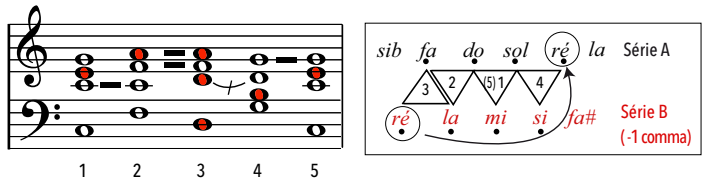


Figure 8.32 Enchaînement des accords parfaits d'intonation pure avec remontée du *ré* sur le 4e accord, afin de rester au diapason de départ. Présentation en musique (les traits indiquent les notes communes conservées), et en diagrammes de quintes.

On trouvera un exemple équivalent dans l'étude d'O. Bettens[49]. L'auteur conclut : « Tout bon chanteur (ou groupe de chanteurs) est amené à se construire un repère diatonique, sorte de canevas, explicite ou implicite, sur lequel il s'appuie pour "s'accorder", c'est-à-dire ajuster son intonation. » Nous avons exposé en détail cette difficulté, car elle est emblématique des questions concrètes que les musiciens sont tenus de résoudre. L'incompatibilité que nous venons d'entendre et le comma qui en résulte peuvent être mis en évidence par le calcul des intervalles (voir annexe B, § 2). Les deux autres commas que l'on rencontre en accordage sont le comma pythagoricien et le comma enharmonique.

Quintes pures et octave : le comma pythagoricien (Son 8.26)

Le comma pythagoricien doit son nom au fait que la quinte pure, de rapport 3/2, est attribuée à Pythagore. Il traduit l'incompatibilité entre quintes pures et octave pure. Partant de do_3, nous accordons une succession d'intervalles purs : quintes ascendantes et quartes descendantes, ce qui génère la succession : *sol, ré, la, mi, si, fa#, do#, sol#, ré#, la#, mi#, si#* (figure 8.33). Si nous assimilons le *si#* à un *do*, un nouveau problème se pose : le *si#* issu de 12 quintes et quartes pures est plus haut que l'octave du *do* de départ (Son **8.26**). L'écart entre ce *si#* (rouge) et le *do* est un comma pythagoricien.

Son 8.26 (36")

Séquence d'accordage démontrant le comma pythagoricien ; orgue électronique *(voir page 441)*

49. *Intonation juste à la Renaissance : idéal ou utopie ?*, voir le site www.virga.org/zarlino/ qui comporte un grand nombre d'exemples sonores et un logiciel libre permettant d'expérimenter soi-même.

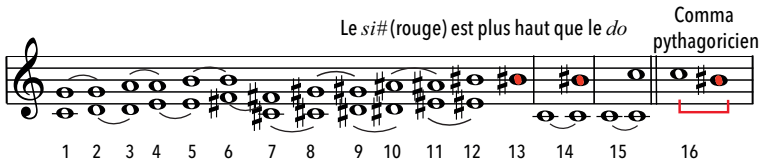


Figure 8.33 Le problème du comma pythagoricien : le *si#* obtenu par la 12e quinte est plus haut que le *do*.

Tierces pures et octave : le comma enharmonique (Son 8.27)

Le troisième comma, dit enharmonique, traduit l'incompatibilité entre tierces pures et octave pure. Le comma enharmonique est l'écart entre le *si#* résultant de l'enchaînement de 3 tierces majeures pures et le *do* de l'octave de départ (figure 8.34). Le comma enharmonique est pratiquement deux fois plus grand que le comma pythagoricien. Voir annexe B, § 2 pour le calcul de ces commas en cents.

Son 8.27 (23")

Séquence d'accordage démontrant le comma enharmonique ; orgue électronique *(voir page 441)*

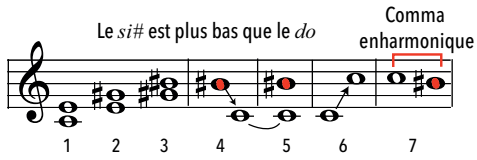


Figure 8.34 Comma enharmonique : le *si#* est plus bas que le *do*.

3.2. La réalisation pratique des systèmes d'accord d'instruments de hauteur fixe

Sur un instrument à sons fixes, les systèmes qui n'utilisent que des intervalles purs imposent de multiplier le nombre de degrés. Ils se distinguent nettement des systèmes qui sont limités à 12 degrés dans une octave pure, que l'on qualifie de tempéraments. Indépendamment des considérations numériques développées par les théoriciens, le succès de l'un ou l'autre de ces systèmes d'accord s'explique prioritairement par la recherche d'une adéquation entre un type de musique et les contraintes instrumentales de sa réalisation sonore.

3.2.1. Les systèmes à intervalles purs : systèmes ouverts

Il peut paraître surprenant, pour les musiciens d'aujourd'hui, que nos prédécesseurs se soient donné la peine de concevoir, construire et jouer des claviers ayant jusqu'à 31 touches dans l'octave. Pourtant, c'est la seule solution pour disposer simultanément des intervalles de quinte et de tierce pures (quinte 3/2, tierce mineure 6/5 et tierce majeure 5/4) sur tous les degrés de l'échelle chromatique. Pour reprendre l'exemple de la suite d'accords présentée sur la figure 8.31, ce type d'instrument offre deux touches différentes pour le *ré* : une pour le *ré* quinte pure du *sol*, et une pour le *ré* tierce mineure pure du *fa*. Il ne faut donc pas imaginer que la multiplication du nombre de touches dans l'octave répondait à une recherche de microtonalité, mais bien à la possibilité d'offrir, sur un instrument polyphonique à sons fixes, les différents ajustements de hauteur permettant à tout instant de rester dans une harmonie d'accords purs. Le rôle important joué par les tierces dans l'harmonie de la Renaissance justifie à lui seul de telles recherches. Un des premiers instruments connus est l'archicembalo de Nicola Vicentino[50].

50. *L'antica musica ridotta alla moderna prattica*, Nicola Vicentino, 1555.

Marin Mersenne explique la disposition de claviers de 17, 19, 27 et 32 touches à l'octave (figure 8.35), mais son texte est peu accessible au lecteur non musicologue en raison des termes en usage pour désigner les sons[51].

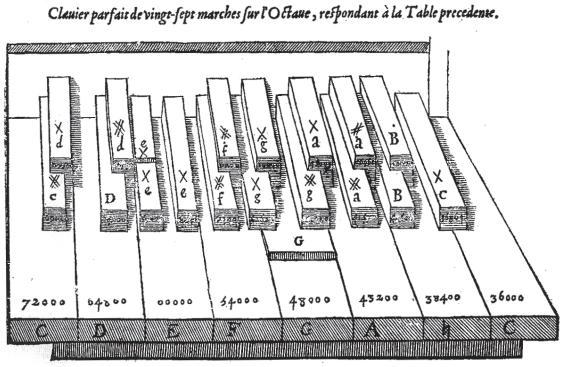


Figure 8.35 Clavier harmonique, *Parfait* de 27 « marches » sur l'octave.

Mersenne, *Harmonie universelle*, Livre sixième des orgues, Proposition XXIII, p. 356 [fac-similé CNRS, 1965].

Le corollaire d'un accord par intervalles purs est l'instauration d'une gamme à degrés inégaux : il existe des tons majeurs et des tons mineurs qui donnent au chromatisme une saveur étonnante.

Pour étudier plus avant cette question, nous engageons vivement le lecteur à se reporter aux nombreux ouvrages existant[52], ainsi qu'aux sites Internet consacrés à la musique de la Renaissance. L'informatique musicale offre aujourd'hui la possibilité d'entendre les madrigaux de Marenzio et les chansons de Claude Lejeune (en particulier la chanson *Qu'est devenu ce bel œil ?*) avec toute la justesse requise.

L'accordage des instruments polyphoniques a aussi trouvé écho auprès de nombreux théoriciens modernes. Helmholtz, qui consacre un important chapitre aux gammes, a fait construire un harmonium à deux claviers, spécialement accordé sur les intervalles « justes » selon sa terminologie.

À la même époque, Bosanquet (1872), puis Fokker (1955)[53] ont réalisé des instruments expérimentaux dont le clavier est plus adapté à la recherche acoustique qu'au jeu musical, mais qui sus-citent un regain d'intérêt dans le courant actuel des musiques microtonales. C'est un sujet inépuisable qui inspire toujours de nombreux ouvrages et pas mal de discussions.

Pour clore ce sujet, nous renvoyons le lecteur au site Internet d'Olivier Bettens qui répond à la question « *Intonation juste à la Renaissance : idéal ou utopie* ? » avec beaucoup de finesse dans une longue conclusion intitulée « L'oreille et la raison ». (http://virga.org/zarlino/index.html)

51. Mersenne, 1636, *op. cit.*, Livre sixième des orgues, Proposition XXII, *Expliquer la science du clavier des orgues...*, p. 349-358.
52. Asselin (1985), Barbour (1972), Legros (1972), Lattard (1988).
53. L'harmonium de Bosanquet est à Londres (tardis.dl.ac.uk/FreeReed/organ_book/node22.html). L'orgue de Fokker a été récemment restauré en Hollande (www.huygens-fokker.org/docs/mm4.html). Voir sur ce site les textes de Pels (1950) et Fokker (1955).

3.2.2. Les systèmes fermés de 12 touches dans l'octave : les tempéraments inégaux

> TEMPÉRAMENT - *Opération par laquelle, au moyen d'une légère altération dans les intervalles, faisant évanouir la différence de deux sons voisins, on les confond en un, qui, sans choquer l'oreille, forme les intervalles respectifs de l'un et de l'autre. Par cette opération on simplifie l'échelle en diminuant le nombre des sons nécessaires. Sans le tempérament, au lieu de douze sons seulement que contient l'octave, il en faudrait plus de soixante pour moduler dans tous les tons.*
>
> Rousseau, J.-J., 1768, *Dictionnaire de musique.*

Comme le formule si finement et si explicitement J.-J. Rousseau, les tempéraments sont des compromis permettant de s'en tenir à douze sons dans une octave pure, sans « choquer l'oreille ».

Aucun des deux intervalles – quinte, tierce – n'étant compatible avec l'octave, il faut procéder à des choix : privilégier un nombre limité d'intervalles purs et altérer les autres afin de boucler l'octave. Le choix peut se porter sur la quinte, ou sur la tierce. On peut aussi altérer tous les intervalles d'une même catégorie de la même quantité (tempéraments égaux) ou selon des règles variables.

Les tempéraments ayant existé sont en très grand nombre et la plupart sont inégaux. Commençons par ceux qui privilégient la quinte pure.

Le tempérament pythagoricien : accord par quintes pures

Un tempérament pythagoricien conserve un maximum de 11 quintes pures. La 12e quinte absorbe l'écart entre le 12e son d'une série de quintes pures et l'octave. Elle est donc plus courte que les autres d'un comma pythagoricien : c'est une quinte « loupée » dite quinte du loup. Classiquement, on place cette mauvaise quinte sur les degrés les moins usités : *si-fa*# (système dit d'Arnaut de Zwolle, vers 1450) ou encore sur *sol*#*-mib*.

La progression par quintes pures ascendantes conduit à réaliser des dièses hauts. À l'inverse, la progression par quintes pures descendantes aboutit à des bémols bas. Il n'existe donc pas d'enharmonie, le choix d'un *do*# ou d'un *réb* devant être décidé à l'avance.

Le tempérament mésotonique : accord par tierces pures et quintes réduites

Le tempérament mésotonique (dit aussi à tons moyens, *meantone*), qui comporte huit tierces majeures pures, témoigne de l'importance prise par les tierces dans la musique du XVIe siècle. Il reste présent, sous différentes variantes, dans le monde de l'orgue jusqu'à la fin du XVIIIe siècle, comme en témoigne l'exemple de la figure 8.36, qui est la méthode habituelle pour réaliser la partition (Dom Bédos, 1766).

Son 8.28 (15")

Quintes réduites, tierce pure

Pour aboutir à une tierce pure *do-mi*, il faut réduire les quatre premières quintes *do-sol, sol-ré, ré-la, la-mi* d'un quart de comma syntonique (Son **8.28**) en distribuant les battements de façon régulière entre les quintes et les quartes, de sorte qu'ils s'accélèrent très légèrement dans le sens ascendant.

Son 8.28 (15")

Quintes réduites, tierce pure

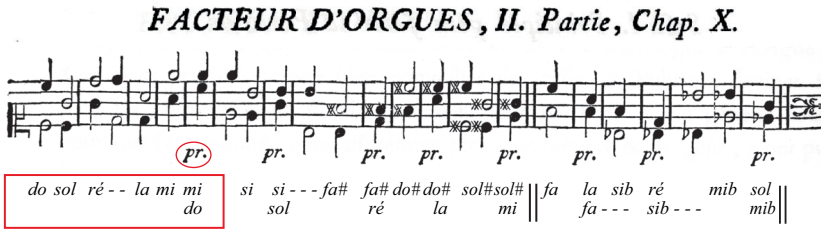


Figure 8.36 Réalisation d'une partition selon l'accord mésotonique.
Le signe *pr.* indique que la tierce doit être sans battement (tierce pure).
Note : cette figure prend en compte les corrections de l'auteur.
Dom Bédos, 1766, *L'art du facteur d'orgues*, p. 432 et 476 [fac-similé L. Laget].

L'opération réussit lorsqu'on aboutit à une tierce *do-mi* sans battement : c'est la première preuve (notée *pr.* sur la figure 8.36). On étend ensuite la partition par quintes ascendantes en vérifiant à chaque fois que les nouvelles tierces sont sans battement (*sol-si ; ré-fa# ; la-do# ; mi-sol#*). Arrivé à ce point on repart d'une nouvelle note, *fa*, par quintes descendantes, en continuant de s'assurer que les tierces sont pures (*fa-la ; sib-ré ; mib-sol*). Toutes les quintes étant fortement réduites, les dièses obtenus par intervalles ascendants sont bas, alors que les bémols, produits par des quintes descendantes, sont hauts, ce qui est à l'opposé du système à quintes pures.

La quinte du loup qui ferme le cycle est très grande : c'est, en toute rigueur, une sixte diminuée, généralement placée sur l'intervalle *sol#-mib*. En bref, huit tierces sont pures; les quatre « mauvaises tierces » restantes étant des quartes diminuées (*do#-fa ; fa#-sib ; sol#-do ; si-mib*). Le tempérament mésotonique que nous venons de décrire représente un point extrême, car les intervalles y sont très différenciés. Il convient à un nombre limité de tonalités.

En modifiant la note de départ, on peut favoriser les tonalités en dièses ou celles en bémols. Le *ré#* et le *mib* étant deux sons distincts, ils n'apparaissent pas simultanément dans une pièce donnée, il faut donc choisir d'accorder l'un ou l'autre des deux sons.

J.-J. Quantz, maître de musique de Frédéric II de Prusse et flûtiste renommé, a poussé le raffinement jusqu'à percer deux trous distincts, l'un pour le *ré#*, l'autre pour le *mib*[54], ce qui l'a conduit à doubler l'unique clé de la flûte traversière de l'époque (voir figure 8.37).

Figure 8.37 Exemple de flûte baroque munie de la double clé (*ré#* et *mib*) préconisée par Quantz (1752).
D'après le *Supplément à l'Encyclopédie* de Diderot, 1777.

54. Voir aussi le portrait du flûtiste Tromlitz, par D. Caffe, tenant entre ses mains une flûte à double clé : Fontana, E., 2015, Portraits of Johann George Tromlitz (1725-1805), *Musiques – Images – Instruments*, vol. 15, CNRS Editions.

Diversité des tempéraments en usage

Il existe une infinité de tempéraments inégaux à tierces pures, qui diffèrent selon le nombre de tierces que l'on souhaite réserver, et selon l'endroit où ces intervalles sont placés dans l'échelle. On recense ainsi des tempéraments français, allemands, italiens, qui conviennent chacun à différentes pièces musicales particulières. Parmi les diverses représentations, nous proposons celle de la figure 8.38, qui permet de visualiser le sens d'altération des intervalles avec un code graphique simple : trait fléché : intervalle pur ; arc de cercle concave : intervalle réduit ; arc de cercle convexe : intervalle agrandi.

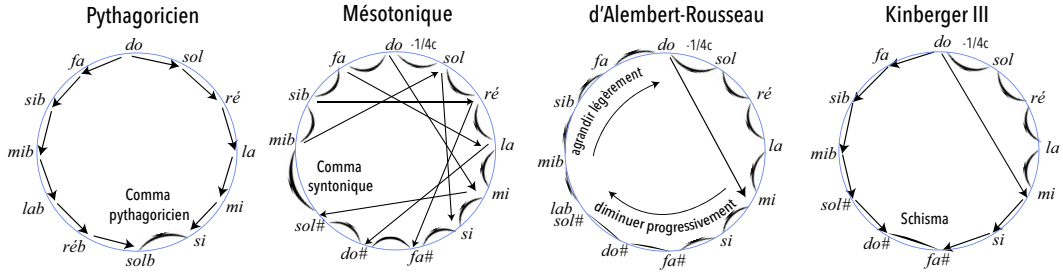


Figure 8.38 Représentation graphique de quatre tempéraments d'après différents auteurs. Certaines quintes sont très peu réduites, d'autres sont même agrandies. La réduction est exprimée en fraction de comma syntonique.

Gammes chromatiques, tempérament égal et mésotonique (voir page 441)

Diapason 440 Hz

Son 8.29a (27″)

Clavecin (égal puis mésotonique)

Son 8.29b (35″)

Cantor (égal puis mésotonique)

Diapason 415 Hz

Son 8.29c (27″)

Clavecin (égal)

Son 8.29d (25″)

Clavecin (mésotonique)

Un chromatisme inégal

L'intérêt porté aux différents systèmes d'accord s'est renouvelé lorsqu'il est redevenu possible de jouer les musiques des XVIe et XVIIe siècles sur les instruments d'origine : clavecin, luth, orgue, dont la richesse harmonique permet de percevoir finement les petites différences entre les intervalles (voir § 2.2.1).

Sensibilisés à l'écoute de ces instruments, les auditeurs ont pu ainsi goûter aussi au plaisir des tempéraments inégaux qui offrent une grande variété d'effets sonores dans le chromatisme (Sons **8.29a**, **b**, **c** et **d** ; figure 8.39) et affectent sensiblement les qualités sonores d'un instrument donné, en particulier le clavecin. Une musique servie par l'instrument et le système d'accordage qui lui conviennent révèle les tensions mélodiques et les épanouissements harmoniques que le compositeur a savamment ménagés dans son écriture.

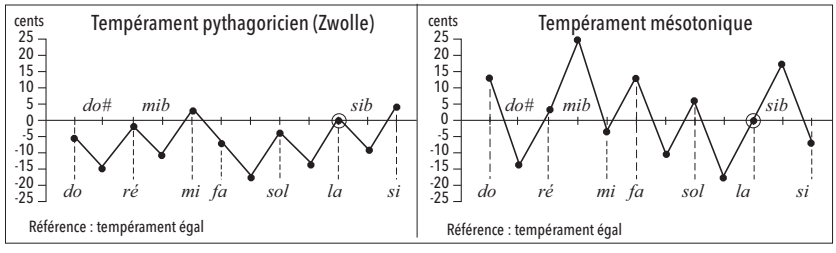


Figure 8.39 Caractéristiques musicales des intervalles inégaux des tempéraments pythagoricien et mésotonique. Écarts en cents par référence au tempérament égal.

D'après Asselin, P.-Y., 1985, figure 129.

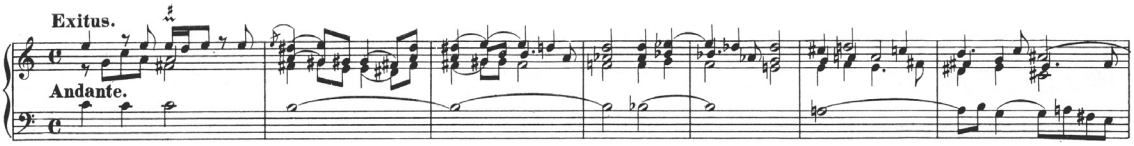


Figure 8.40 Extrait du *Kleines harmonisches Labyrinth*, œuvre attribuée à J.-S. Bach (BWV 591).
Source : IMSLP.

Il faut bien admettre que la pratique personnelle de l'accordage d'un instrument conditionne en grande partie l'accès à ce mode sonore.

Musique (Cantor)
(voir page 441)

C'est pourquoi nous avons retenu un exemple musical particulièrement contrasté qui permet de comparer le même fragment joué sur un instrument électronique accordé tout d'abord au tempérament égal (Son **8.30**), puis au tempérament mésotonique (Son **8.31**). La pièce choisie[55] (figure 8.40) est riche en chromatismes. Elle comporte des enharmonies et d'incessantes modulations pour lesquelles l'incidence de l'accordage est particulièrement frappante.

Son 8.30 (34")
Tempérament égal

Les exemples musicaux enregistrés sur orgue électronique sont un pis-aller qui ne remplacera jamais le jeu sur instrument réel dont le timbre se modifie selon le type d'accordage.

Son 8.31 (36")
Tempérament mésotonique

Pour entrer dans le monde des tempéraments et développer une sensibilité à l'écoute de leur réalisation, le lecteur doit se procurer les exemples sonores réalisés par Pierre-Yves Asselin à l'orgue et au clavecin[56] (Asselin P.-A. et Legaillard Y., 1985). Les exemples choisis permettent d'entendre des extraits d'œuvres caractéristiques des différents types d'accord présentés. Le même extrait musical est joué successivement sur l'instrument accordé au tempérament historique requis pour une pièce donnée, puis au tempérament égal qui sert de référence.

Ces exemples montrent bien qu'un système d'accord est un triple compromis : acoustique (il dépend des contraintes physiques et de l'instrument), musical (il dépend de l'écriture musicale) et esthétique (il dépend de la culture et des goûts du moment).

Dans un tempérament inégal, les tonalités sont reconnaissables aux changements de tension sur certains accords, ce qui a donné lieu à plusieurs descriptions métaphoriques, dont celles de Mattheson[57] et de Marc-Antoine Charpentier (figure 8.41).

Do M	Gai et guerrier
do m	Obscur et triste
Ré M	Joyeux et très guerrier
ré m	Grave et dévôt
Mib M	Cruel et dur
mib m	Horrible, affreux
Mi M	Querelleux, criard
mi m	"Efféminé", amoureux et plaintif
Fa M	Furieux et emporté
fa m	Obscur et plaintif
Sol M	Doucement joyeux
sol m	Sérieux et magnifique
La M	Joyeux et champêtre
la m	Tendre et plaintif
Sib M	Magnifique et joyeux
sib m	Obscur et terrible
Si M	Dur et plaintif
si m	Solitaire et mélancolique

M.A. Charpentier (1636-1704),
Règles de composition - Paris, 1690

Figure 8.41 Caractères des tonalités selon Marc-Antoine Charpentier.
Source : Pierre-Alain Clerc, *Discours sur la rhétorique musicale* (article téléchargeable à partir du titre).

55. Le tempérament mésotonique a été choisi pour accentuer les oppositions sonores, il n'a pas de justification historique pour J.-S. Bach.
56. En attente d'une réédition aux éditions Lemoine, quelques extraits sont consultables dans les archives du LAM à l'adresse http://telemeta.lam.jussieu.fr
57. J. Mattheson, *Das neu-eröffnete Orchestre*, 1713. M.-A. Charpentier, *Règles de composition*, 1690. Voir Pierre-Alain Clerc, *op.cit.* p. 46-48.

3.2.3. Le tempérament égal à 12 sons dans l'octave

Le tempérament égal à 12 sons dans l'octave est connu depuis longtemps. Marin Mersenne le décrit très précisément pour l'accord des instruments frettés, luth et viole (figure 8.42). Parallèlement aux méthodes géométriques et aux calculs des théoriciens, les musiciens adoptaient une règle simple consistant à appliquer la proportion 18/17[58] dont la valeur (1,058) est très proche de celle de la racine douzième de 2 (1,059). Bien qu'un peu plus faible, cette proportion convient en pratique, car le fait d'appuyer sur la frette produit une légère augmentation de tension et une élévation du son.

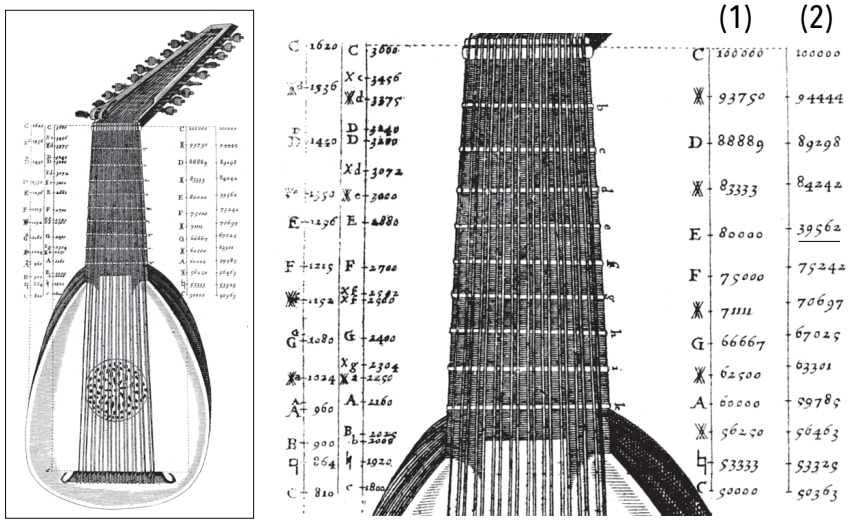


Figure 8.42 Le luth et le frettage du manche. Sur la vue agrandie on voit : à gauche du manche, le frettage avec intervalles purs : respectivement 13, 19 sons à l'octave ; à droite du manche, le frettage au tempérament égal : (1) théorique, (2) réalisé sur le luth.

Mersenne, 1636, *Harmonie universelle,* Livre second des instruments à cordes, Proposition I, p. 46 [fac-similé CNRS,1965].

Le tempérament égal s'est généralisé au XIXe siècle pour répondre aux besoins de l'écriture musicale (modulations, chromatisme). Il a en outre bénéficié du *flou* du piano (battements des tricordes et inharmonicité du son), de la dispersion de justesse produite par les grandes masses instrumentales et de l'usage généralisé du vibrato. Les résistances, qui ont été nombreuses, réapparaissent à propos du renouveau des instruments de tradition ancienne comme le clavecin.

Le tempérament égal est devenu la base de référence théorique de la musique occidentale et s'est répandu assez largement. Si le choix d'un accord est affaire de goût, les considérations pratiques l'emportent parfois, au détriment de la qualité sonore, par exemple lorsqu'il est nécessaire de transposer rapidement d'un demi-ton au cours d'un concert. Exception faite des instruments électroniques, l'accord en tempérament égal est rarement réalisé de façon rigoureuse, car il est difficile à contrôler. Il faut en théorie réduire chaque quinte de 1/12 de comma pythagoricien. Au milieu du clavier les battements sont lents et doivent progresser très régulièrement, de sorte que la quinte à l'octave supérieure de celle du départ batte exactement deux fois plus vite.

58. Mark Lindley résume ainsi l'opération : « diviser la corde en 18 parties; placer la 1re frette à la première division. Diviser la portion restante en 18 parties, placer la 2e frette à la première division ; diviser la portion restante en 18 parties, et ainsi de suite jusqu'à la 12e frette. » *op. cit.* p. 198.

On présente généralement les tempéraments dans une perspective historique liée à l'évolution de l'écriture musicale. Cependant, les contraintes instrumentales sont au moins aussi importantes que les contraintes musicales. En remarquant qu'au XVIIe siècle, les instruments à frettes étaient au tempérament égal, alors que le clavecin et l'orgue étaient manifestement accordés en mésotonique, on est tenté de se rallier à l'opinion des auteurs qui avancent l'idée d'une coexistence pratique de différents systèmes d'accord (Bougeret, 1982). Cette vue relativise aussi l'idée, toujours tenace, selon laquelle le tempérament égal, qui s'est imposé au XXe siècle, témoignerait d'un progrès, et donc serait la solution idéale de l'accordage, alors qu'il s'agit seulement du meilleur (ou du moins mauvais) des compromis.

3.3. Accordage et instrument : de la théorie à la réalisation pratique

3.3.1. Le son des instruments réels

Je n'entends pas les battements... Je n'entends rien... !

Asselin, P.-Y., 1985, p. 21.

Son 8.32 (12")

Clavecin ; sons isolés A et B

Son 8.33 (6")

Unisson accordé : A et B à la même fréquence

Son 8.34 (5")

Unisson faux : A et B désaccordés

Telles sont les plaintes des débutants s'initiant à l'accordage du clavecin, et tout apprenti accordeur est passé par ce stade. Une fois les notions théoriques assimilées, la confrontation avec les sons réels est déconcertante. En rapportant ces propos Pierre-Yves Asselin ajoute : « l'oreille humaine entend **tout** et c'est bien là le problème car elle doit apprendre à n'entendre que des phénomènes isolés pour accorder ». Il signifie par là qu'un débutant ne parvient pas à sélectionner les phénomènes sur lesquels il faut s'appuyer pour régler l'accord. S'initier à l'accordage nécessite une éducation de l'écoute pour apprendre à négliger certaines fluctuations normales du son, afin de se concentrer sur les indices du son qui sont pertinents pour la justesse. Commençons par un exemple très simple : la mise à l'unisson de deux notes d'un clavecin.

L'expérience suivante pose clairement les problèmes rencontrés avec les sons naturels, et permet de comprendre le désarroi d'un débutant qui commence par ajuster l'unisson entre deux sons. Soient deux sons A et B (si_2) produits par deux cordes différentes du même clavecin (Son **8.32**). La hauteur est bien stable. Ils ont été accordés pour former l'unisson (Son **8.33**). Avant l'accord, les deux sons formaient de violents battements (Son **8.34**). L'analyse spectrographique de la figure 8.43 surprend à première vue : il semble qu'il y ait autant de battements sur les sons isolés, sur l'unisson accordé et sur l'unisson faux.

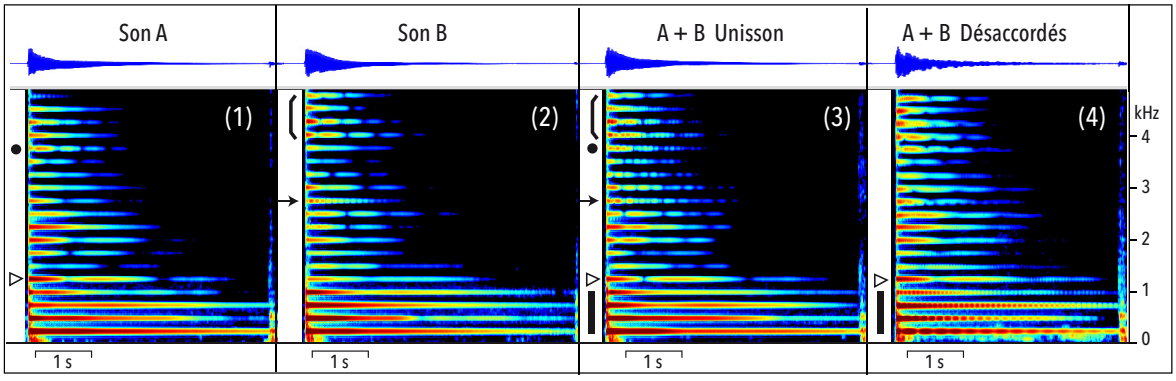


Figure 8.43 L'analyse spectrographique des sons isolés A et B révèle des instabilités (battements de partiels) qui se retrouvent dans le spectre de l'unisson A + B et troublent la perception d'un apprenti accordeur.

Examinons séparément les composantes graves du son et la partie aiguë du spectre. En comparant les quatre premières composantes graves de la figure, on voit que les battements visibles en (4) ne sont pas dans (3). Quant aux fluctuations d'amplitudes localisées dans le haut du spectre, elles n'ont pas d'incidence directe sur l'appréciation de justesse mais participent à la sensation globale de qualité du son. Certains « accidents », particuliers aux notes isolées, se retrouvent dans l'unisson (voir les signes sur la figure 8.43 qui indiquent, pour le son A : les battement des partiels 5 et 15 ; pour le son B : les battements du partiel 11 et la zone aiguë autour de 4 kHz). Cette analyse montre les limites d'une approche théorique purement numérique. Elle peut aussi expliquer les piètres résultats des accordeurs électroniques qui mesurent la seule fréquence fondamentale alors que l'écoute humaine intègre l'ensemble des qualités du son, et en particulier la qualité sonore d'un agrégat, pour décider si l'accord est satisfaisant ou non. Le lecteur expérimenté écoutera avec intérêt le Son **8.37** capté au cours de l'accordage d'un clavecin par un accordeur professionnel[59], dont l'analyse détaillée est donnée page 438.

3.3.2. Quelques problèmes relatifs à l'accordage du piano

Nous n'avons pas traité du piano pour deux raisons. Tout d'abord, parce que l'ajustement et la stabilisation des chevilles d'accord nécessite un apprentissage particulier, ensuite, parce que les phénomènes acoustiques y sont beaucoup plus complexes[60]. Le son est inharmonique (voir chapitre 6, § 3.2) et il y a plusieurs cordes par note (excepté dans le grave). Un bon accordeur ménage un infime battement entre ces cordes, ce qui confère « de la vie » au son, mais contribue au flou perceptif. Plus encore que pour le clavecin, l'accordage d'un piano repose sur l'écoute de la sonorité des accords, sur l'impression globale de fusion des sons et de leurs mouvements, plus que sur une arithmétique compliquée. On pourra écouter (Son **8.35**) et comparer le spectre de deux notes isolées de piano (figure 8.44) avec ceux du clavecin vus précédemment, et remarquer que même les composantes graves sont instables en amplitude.

Son 8.35 (15")

Les deux sons de piano de la figure 8.44

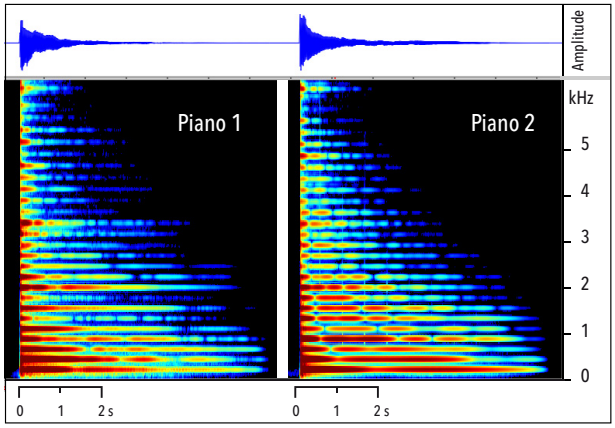


Figure 8.44 Analyse spectrographique de deux sons de piano où l'on voit qu'une note isolée est animée de modulations d'amplitude complexes.

Source des sons : piano 1 (McGill), piano 2 (Iowa).

59. Accordeur des clavecins au CNSMP. Voir Marandas & coll., 1998, *op. cit.*, piste n° 14 du CD.
60. De nombreux ouvrages y sont consacrés, voir en particulier J. Lattard, 1988 et 1997.

Pour toutes ces raisons, le relevé d'accord d'un piano réglé sur le tempérament égal s'écarte franchement de la référence théorique (figure 8.45).

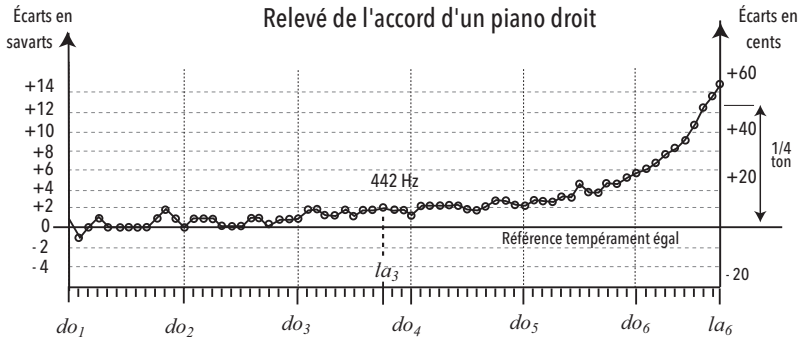


Figure 8.45 Relevé de l'accord des cinq octaves supérieures d'un piano droit. L'axe horizontal est la référence du tempérament égal, la_3 = 440 Hz.

Par rapport au tempérament égal idéal, la courbe monte du grave à l'aigu, car les octaves sont faiblement mais régulièrement agrandies. Dans la dernière octave, cet agrandissement est fortement accentué pour répondre aux particularités de l'oreille (voir chapitre 3, § 2.7.3).

Nous proposons (Son **8.36)** l'écoute d'un exemple sonore emprunté à Taro Mori (2000), qui permet d'apprécier l'effet d'un agrandissement de l'octave entre deux sons de piano numérisés dont on règle l'inharmonicité (voir les commentaires détaillés dans la légende du son, page 442).

Son 8.36 (23")

Effet de l'inharmonicité sur la sensation de justesse d'un intervalle d'octave au piano
(voir page 441)

3.3.3. La pratique de l'accord sur divers instruments

Chaque type d'instrument pose des problèmes d'accordage spécifiques. Tout d'abord en raison de la structure acoustique du son : plus ou moins riche en harmoniques, quasi harmonique ou franchement inharmonique (carillons). Ensuite parce qu'il existe une grande diversité de combinaisons adaptées au style de musique. Le cas le plus simple est celui des instruments polyphoniques dans lesquels chaque son est produit par un système vibrant indépendant. L'orgue en est un exemple emblématique. Toutefois, les actions à effectuer pour accorder un tuyau d'orgue ayant une incidence non négligeable sur le qualité du son, il s'y rencontre aussi des compromis à établir.

Parmi les instruments non entretenus, la harpe à pédales a toujours posé des problèmes particuliers. C'est un instrument de grande étendue qu'il faut accorder avant chaque concert. La faible tenue du son, le système mécanique de raccourcissement des cordes pour produire les sons altérés (dièses et bémols) et surtout l'étendue du répertoire avec orchestre contraignent pratiquement les harpistes à adopter le tempérament égal. Aujourd'hui, les accordeurs électroniques ont supplanté l'accordage à l'oreille, essentiellement pour des raisons de gain de temps.

La famille des instruments à cordes et à frettes pose des problèmes presque insolubles : il faut concilier des cordes à vide, et différentes occurrences de sons se trouvant à plusieurs endroits du manche. Chaque musicien développe sa stratégie d'accord, stratégie qu'il doit sans cesse adapter au renouvellement des cordes. Pour la partie historique, nous renvoyons le lecteur à l'ouvrage de Mark Lindley, *Luths, viols & temperaments*, 1984.

3.3.4. Harmonie et mélodie : deux justesses ?

La justesse, écart par rapport à une référence en mémoire, est une notion culturelle très relative. Mais lorsque deux musiciens jouent ensemble, l'un sur un instrument à son fixe et l'autre sur un instrument à sons variables, des tensions peuvent se manifester, puisque les références de justesse ne sont pas les mêmes pour l'harmonie et pour la mélodie.

Prenons l'exemple d'une sonate pour violon et clavecin. Si les tempéraments à tierces pures sont de loin préférables pour la beauté des accords et pour la plénitude sonore du clavecin, ils ont pour inconvénient de générer une échelle mélodique étrange aux oreilles d'un violoniste porté à l'expressivité de la mélodie. Les tierces majeures pures, étrangement basses, et les altérations inversées (dièses bas et bémols hauts) sont en contradiction avec la justesse mélodique qu'il recherche et qu'il a travaillée.

Pourtant, l'aventure mérite d'être tentée, car, après quelque temps, les deux musiciens adaptent leur jeu en usant des ressources de l'ornementation pour esquiver les rencontres malheureuses, en particulier les doublures de tierces. La musique se développe alors dans une dimension esthétique nouvelle, en bénéficiant de la plénitude sonore qu'offre l'accordage adapté à l'instrument polyphonique.

4. Le diapason et l'oreille dite absolue

Accorder un instrument nécessite un étalon, une référence. Un musicien soliste de tradition orale ne s'en soucie guère : il fait sonner son instrument au mieux. Mais, s'il veut joindre la voix à l'instrument ou s'il joue avec un hautboïste qui ne peut guère modifier la hauteur des sons, il devra en tenir compte pour accorder son instrument.

Dès le XVIIe siècle, les musiciens se sont préoccupés de trouver un étalon de référence pour la hauteur des sons[61]. L'orgue a longtemps joué ce rôle, car la note la plus grave d'un tuyau de flûte dépend directement de la longueur du tuyau. Cependant, les unités de mesure comme le pied[62] variaient selon les régions et les pays, ce qui entraînait une grande dispersion des fréquences d'accord. Les flûtistes qui voyageaient devaient alors posséder plusieurs « corps de rechange » pour adapter leur instrument à la diversité des fréquences d'accord selon les provinces. L'étude de textes anciens (Praetorius, 1618 pour les orgues), ainsi que le relevé dimensionnel des flûtes et hautbois déposés dans les musées, ont permis de rendre compte de cette variabilité (Haynes, 2002 ; Leipp & Castellengo, 1977).

61. Voir Théodora Psychoyou, 2003, *op. cit.*, tome 1, p. 180.
62. Unité de mesure du système duodécimal (à base 12), en usage au XVIIe siècle. Un pied vaut 12 pouces, un pouce vaut 12 lignes, une ligne vaut 12 grains.

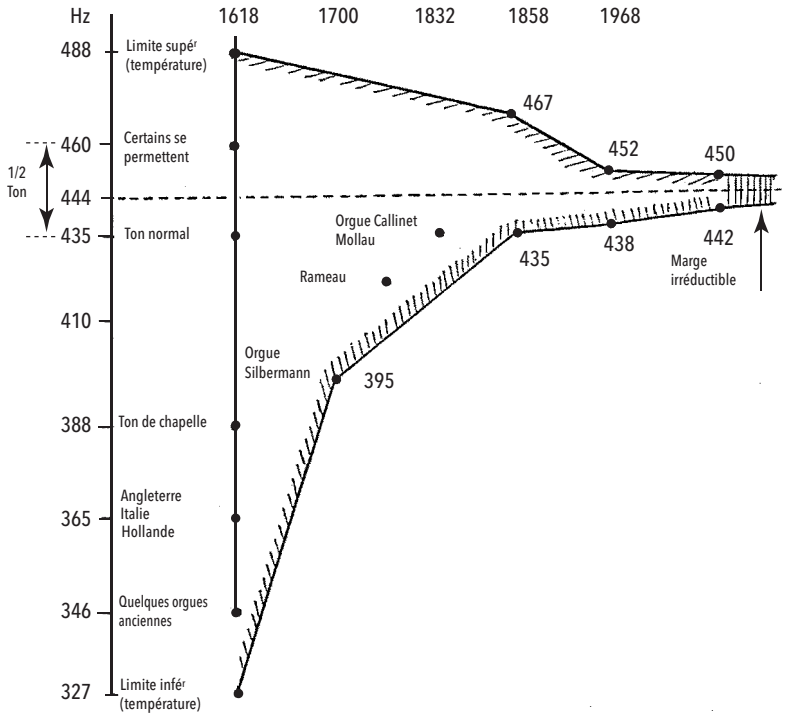


Figure 8.46 Dispersion du diapason, de 1618, époque de Praetorius, à nos jours.

Leipp E. & Castellengo M., 1977, figure 1.

La figure 8.46 donne une idée de la variabilité du diapason d'un lieu à un autre et au cours des âges. Au fur et à mesure qu'on se rapproche de l'époque moderne, les unités de mesure se normalisent et, dès 1859, les fabricants d'instruments de musique doivent respecter la norme d'accord édictée par la commission Lissajous-Halévy[63]. D'autres normes européennes et mondiales suivront. La dernière en date[64] pour l'Europe recommande de normaliser la fréquence d'accord à 440 Hz à 20 °C. Parallèlement, et pour des raisons de pratique instrumentale, un *diapason ancien* à 415 Hz s'est implicitement généralisé[65].

Le diapason est l'objet de nombreuses discussions de la part des musiciens et donne lieu le plus souvent à des discours non fondés. C'est un nombre fixant la fréquence. Or, comme on le sait, la sensation de hauteur ne dépend pas seulement de la fréquence d'un son, mais aussi de son spectre. Par ailleurs, les conditions de jeu – température, hygrométrie – ont des effets différents sur l'accord des instruments en cours de jeu : certains montent, d'autres restent stables au cours d'une exécution orchestrale. Les critiques les plus acerbes émanent le plus souvent de chanteurs qui donnent foi aux rumeurs affirmant une montée permanente, mais aussi de musiciens ayant acquis l'oreille absolue, et qui souffrent des plus infimes décalages par rapport à la référence standard.

63. Après une enquête européenne le diapason est fixé à 435 Hz et le Conservatoire de Paris se dote d'une cloche (son de hauteur ambiguë !) sonnant le *la* officiel. Voir Leipp & coll. *Bulletin du GAM* n° 88.
64. 16e résolution du Conseil de l'Europe, 1971.
65. Le plus répandu est la_3 = 415 Hz, soit un demi-ton en dessous. Certains musiciens utilisent aussi un la_3 à 392 Hz.

Quelques remarques sur l'oreille absolue

Il faut tout d'abord relativiser la dénomination « d'oreille absolue », que l'on doit plutôt attribuer à une excellente mémoire des hauteurs sonores acquise dans la petite enfance, particulièrement entre cinq et sept ans. L'expérience personnelle de plusieurs musiciens, de même que les récentes études qui s'appuient sur l'écoute de sons réels instrumentaux plutôt que de sons synthétiques, confirment que cette mémoire concerne prioritairement un instrument donné, celui que l'on pratique. L'apprentissage est plus aisé avec un instrument de hauteur fixe (piano) et se stabilise plus rapidement lorsque le son de l'instrument, la voix et le nom de la note sont associés. C'est donc prioritairement une mémoire liée au matériau sonore, que plus d'une personne sans éducation musicale particulière peut posséder, sans en avoir conscience. Ce n'est que par l'association répétée des noms de notes avec des sons de hauteur précise, et de sources différentes, que l'écoute et la mémorisation s'organisent pour généraliser cette compétence à tous les types d'instruments, au fur et à mesure de l'élargissement de la pratique musicale. Par bonheur, l'oreille absolue est tolérante et, compte tenu de la catégorisation de l'échelle musicale selon 12 notes, elle admet des écarts de l'ordre du quart de ton. Cependant, la pratique d'un diapason *ancien*, plus bas d'un demi-ton, crée de véritables problèmes pour les oreilles des musiciens ayant acquis cette capacité. Ceux qui peuvent s'adapter remarquent que le changement de référence, qu'ils finissent par accepter, s'appuie sur la reconnaissance des timbres identitaires : la note donnée par la touche la_3 d'un clavecin à 415 Hz, bien qu'étant un $sol\#_3$ pour l'oreille absolue, finit par être entendue comme un « la_3 », sans pour autant perdre les références des hauteurs de notes sur les autres instruments au diapason 440 Hz. Mais c'est une première brèche dans la fiabilité de cette acquisition. D'autres désarrois assaillent le musicien sûr de son oreille : les sons synthétiques qui lui font perdre son assurance, les moments de fatigue qui peuvent provoquer des décalages, et enfin le vieillissement qui fragilise toutes les fonctions cognitives. La plupart des musiciens se plaignent d'entendre trop bas lorsqu'ils avancent en âge, ce qui par ailleurs continue d'alimenter les opinions sur la hausse du diapason.

L'oreille absolue a suscité un grand nombre de travaux scientifiques, car ce phénomène intrigue vivement ceux qui n'y ont pas accès. Les expérimentations conduites avec des sons sinusoïdaux donnent des résultats confus en raison même de la dissociation hauteur/timbre. Les recherches les plus récentes, s'intéressant au développement de cette compétence chez les musiciens, intègrent dans leur protocole les instruments et leur timbre, l'âge d'imprégnation, la différenciation entre hauteur tonale et hauteur spectrale, la mémoire proprioceptive, la pédagogie musicale, et mettent à profit les techniques des neurosciences pour tenter de débusquer les zones corticales en rapport avec cette aptitude[66].

Il existerait, semble-t-il, autant de types d'oreille absolue que de pratiques musicales[67]. À votre tour, cher lecteur musicien de nous faire part de votre expérience, afin de contribuer à enrichir une question trop souvent abandonnée à l'investigation des seuls scientifiques fascinés et intrigués, à juste titre, par une compétence qui échappe à la mesure.

66. Voir les textes de Miyazaki (1988 & 2004), Takeuchi (1993), Zatorre (2003), Levitin (2005) et, en français, Vangenot (2005).
67. Voir Bachem (1937).

Accordage d'une quinte pure au clavecin (la_3, 415 Hz)

Son 8.37 (47")

Séquence d'accordage de la quinte fa_2-do_3.

Au cours de la séquence d'accordage, qui dure un peu moins d'une minute, l'accordeur répète 14 fois la quinte, assez régulièrement, tout en modifiant la fréquence de la note supérieure do_3. Voici l'analyse du processus d'accordage (figure 8.47).

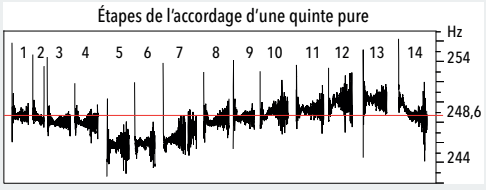


Figure 8.47 Évolution de la fréquence de la note supérieure pendant l'accordage de la quinte. L'analyse est faite avec Praat après filtrage de la fréquence fondamentale du do_3. L'horizontale indique la fréquence cible (248,6 Hz).

De **1 à 4**, l'intervalle bat lentement. L'accordeur abaisse légèrement le do_3, puis franchement en **5.** L'intervalle, trop court, bat rapidement. De **6 à 8,** l'accordeur opère la remontée du do_3 : les battements ralentissent. De **10 à 11,** il dépasse l'intervalle pur. En **12 et 13,** le do_3 est trop haut : l'intervalle, agrandi, bat nettement.

À la dernière quinte, **14,** l'accordeur descend rapidement la note supérieure pendant l'extinction et la cale au tout dernier instant. L'opération n'est pas facile à saisir par une oreille naïve qui continue à entendre « un son qui bouge », car l'instrument possède, comme nous l'avons vu, des fluctuations intrinsèques.

Examinons plus en détail l'accord de la dernière quinte sur la figure 8.48.

Le curseur a été placé sur le sonagramme au moment où le bon point d'accord est atteint et on observe un phénomène intrigant.

Son 8.38 (8")

Ajustement de la dernière quinte (n° 14)

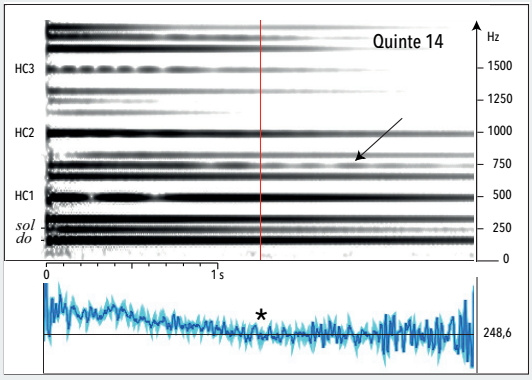


Figure 8.48 Analyse spectrale de la quinte 14 et mesure de la variation de fréquence du fondamental du do_3 : courbe à la partie inférieure. Le curseur désigne le moment où le do_3 est à la bonne fréquence. Les battements sur les harmoniques communs HC1, HC3 ont ralenti puis disparu, mais d'autres se manifestent (flèche) qu'il faut apprendre à ignorer.

Les battements sur les harmoniques communs, indiqués par HC (1, 2, 3), ont disparu, mais de nouveaux battements apparaissent vers 750 Hz (flèche). Il faut les ignorer, et focaliser son écoute uniquement sur les battements des harmoniques communs. L'écoute de l'accordeur est bien une écoute très experte ! Cette expérience justifie aussi la recommandation donnée de modifier l'accord d'un son pendant l'extinction – et non entre deux sons – afin de repérer très précisément les changements qui ne sont dus qu'aux variations de la fréquence d'accord[1].

1. Nous avons remarqué, en comparant les analyses des voies droite et gauche, que les battements n'étaient pas synchrones. Certains accordeurs font de petits mouvements de tête en cours d'accordage, peut-être pour trouver la meilleure position d'écoute binaurale.

5. Les sons du chapitre 8

5.1. Les intervalles entre sons successifs (mélodie)

Son 8.1 – Intervalle de grandeur variable. Le Son 8.1a donne à entendre deux sons de clavecin formant un intervalle de seconde majeure dont la valeur croît à chaque présentation. Il y a neuf intervalles en tout. L'intervalle initial do_3-$ré_3$ et l'intervalle final do_3-mib_3 sont respectivement un ton et une tierce mineure du tempérament égal. Chaque auditeur peut analyser ce qu'il perçoit et noter à partir de quel numéro il bascule du demi-ton à la tierce mineure. Les réponses varient selon les auditeurs. Chaque étape est de 12,5 cents. Le son 8.1b propose la même expérience dans l'ordre décroissant. Le basculement se produit-il au même point ? [M. C.]

Son 8.2 – Exemple d'un chant traditionnel dans lequel on entend un groupe de femmes auquel se joignent périodiquement les hommes qui chantent à l'octave inférieure. On remarquera qu'à l'exception du *fa* aigu et du *sib* medium bien stables, la mélodie se développe essentiellement par glissandos. *Music from Ethiopia*, plage 28, Chant de mariage érythréen, J. Jenkins. [Tangent Records]

Son 8.3 – Chant mongol de syle khoomij. *Inédit Mongolie*, plage 6, Maison des Cultures du Monde, 1989, Auvidis W 260009 [T. Ganbold]

Son 8.4 – Début du chant mongol précédent lu au ralenti, durée 14''. La mesure (avec Praat) des sons 1 à 5 confirme que 2 et 4 sont bien positionnés autour de l'harmonique 7. [M. C.]

Son 8.5 – Solo de guimbarde du Kazakhstan. *The Silk Road : a Musical Caravan*, CD 2, n° 15, SFW CD 40438. Document Jean During, 2001. [Edil Huseinov]

Son 8.6 – Arc musical ngbaka et chant. République centrafricaine, musicien Nicolas Masémokobo. Document Simha Arom, 1967. [Archives LAM]

Son 8.7 – Chant de Côte d'Ivoire. Deux fillettes baoulé (5 et 7ans) alternent avec un chœur de femmes. *Les Voix du monde*, CD 3-13 ; CNRS-Le Chant du monde. Document Hugo Zemp. [CREM]

Son 8.8 – Extrait (18 à 23 '') du chant précédent analysé au sonagraphe. [*Ibid.*]

Son 8.9 – Séparation des voix des deux fillettes. On entend l'extraction par Audiosculpt des voix isolées de la 1re et de la 2e tierce, puis l'intervalle de triton entre le son le plus grave et le son le plus aigu. *Ibid.* [M. C.]

Son 8.10 – Exemple réalisé au clavecin pour apprécier la différence entre les deux tons de la gamme diatonique construite sur les intervalles de la série harmonique. On entend successivement le ton mineur do_3-$ré_3$ (rapport 10/9); puis le ton majeur do_3-$ré_3$ (rapport 9/8). Dans une séquence mélodique, et bien qu'il s'agisse de sons stables, cette différence est tout juste discriminable. [M. C.]

Son 8.11 – Ecart entre les deux $ré_3$ de l'exemple précédent (le premier son est plus bas). Cet intervalle vaut 21,5 cents. [M. C.]

Son 8.12 – Mélodie jouée à la flûte traversière solo (flûte Boehm) par Pierre Séchet, 1977. Haydn, *Londoner trio n° 1*. [Archives LAM]

5.2. Les intervalles entre sons simultanés

Son 8.13 – Battements de deux diapasons à fourche que l'on peut considérer ici comme deux sons simples (purs). On entend successivement le diapason A (440 Hz) puis le diapason B légèrement plus bas. Peu de personnes perçoivent une différence entre les deux sons. Remarque : la frappe de chaque diapason, dont

l'intensité est largement supérieure à celle du son tenu, a été éliminée au montage. [M. C.]

Son 8.14 – Même séquence que précédemment suivie du son des deux diapasons posés ensemble sur la table d'harmonie. Leurs ondes sonores se combinent en produisant 4 battements par seconde. On en déduit que le diapason B est à 436 Hz. [M. C.]

Son 8.15 – Battements de deux sons complexes synthétiques riches en harmoniques, de fréquence 523 Hz et 527 Hz. La rapidité des variations d'amplitude croît avec le rang de l'harmonique (voir figure 8.20). Cependant, on entend seulement un battement de 4 Hz, car tous les harmoniques sont synchronisés sur la variation d'amplitude du fondamental (ou harmonique 1). Synthèse numérique. [M. C.]

Son 8.16 – Perception de battements très lents. Deux sons voisins de 200 Hz dont l'écart en fréquence est très faible (1 cent) produisent un battement dont la période est comprise entre 8 et 9 s. Cette durée dépasse l'empan temporel d'appréciation des rythmes. L'attention est plutôt attirée par la succession des annulations d'harmoniques qui parcourt le spectre dans un mouvement ascendant et descendant : ce phénomène est appelé « phasing ». Synthèse numérique. [M. C.]

Son 8.17 – Comparaison d'une mélodie chromatique (ascendante-descendante) jouée avec des sons sinusoïdaux puis avec des sons riches en harmoniques. Son 8.17a : curieusement, avec les sons sinusoïdaux, c'est à peine si on entend que la mélodie est jouée en quintes parallèles! Son 8.17b : à l'écoute de la même séquence jouée avec des sons riches en harmoniques on découvre que la première quinte bat lentement (2,75 battements par seconde) et que la deuxième est très désaccordée (battements très rapides, 26 battements par seconde. Exemple produit avec un instrument électronique. [M. C., Cantor]

Son 8.18 – Intervalle de quinte dont on entend successivement la note supérieure seule, si_3, puis le son inférieur mi_3. Des battements lents se produisent sur l'harmonique commun, si_4. On entend ensuite la même quinte sans battement : les fréquences des deux sons sont exactement dans le rapport de fréquence 3/2. [M. C., Cantor]

Son 8.19 – Deux exemples de sons multiphoniques. (1) flûte traversière (Fl 8) ; (2) hautbois (Hb 3). [M. C.]

Son 8.20 – Extraction par filtrage des trois composantes principales de l'accord produit par chaque son multiphonique : pour Fl8, composantes n°s 3, 4, 7 ; pour Hb3, composantes n°s 3, 7, 10. Filtrage Audiosculpt. [M. C.]

Son 8.21 – Séquence synthétisée donnant à entendre un groupe de cinq sons disposés selon trois présentations différentes. Tout d'abord en succession, puis mélangés dans un accord crescendo-decrescendo, finalement percutés à la façon d'un gong. Jean-Claude Risset, début de *Mutations.* [J.-C. Risset, GRM]

5.3. Accordage des instruments polyphoniques à sons fixes

Son 8.22 – Démonstration du comma syntonique. L'accordage successif en quintes pures de la séquence *do, sol, ré, la, mi*, aboutit à une tierce, *do-mi,* dite pythagoricienne. Celle-ci, animée de battements rapides, est plus grande que la tierce pure. La tierce pure est jouée ensuite, suivie de la tierce pythagoricienne. On entend finalement les deux *mi* distants d'un comma syntonique. Exemple réalisé avec l'orgue électronique du LAM. [M. C., Cantor]

Son 8.23 – Clavecin. Séquence de cinq accords parfaits, accordés en intervalles purs. L'accordage est effectué de telle sorte que la note commune à deux accords successifs reste à la même fréquence. Le clavecin est ajusté entre les accords 3 et 4 (voir figure 8.31). P.-Y. Asselin, 1984. [Archives LAM]

Son 8.24 – L'exemple précédent, répété cinq fois de suite, aboutit à une baisse du diapason de l'ordre du demi-ton. Après l'annonce « accord du départ », on entend le premier accord au diapason initial, ce qui permet d'apprécier la dérive du diapason. Pour réaliser cet exemple, il a fallu réaccorder le clavecin à chaque reprise de la séquence. Document P.-Y. Asselin, 1984. [Archives LAM]

Son 8.25 – Séquence des accords parfaits de l'exemple précédent chanté par un ensemble de solistes professionnels. Ceux-ci rajustent subtilement la hauteur du *ré* dans la transition du 3e au 4e accord de façon à terminer, sur le dernier accord, au même diapason que celui du début. Document P.-Y. Asselin, 1984. [Archives LAM]

Son 8.26 – Démonstration du comma pythagoricien. Partant du do_3, la progression de l'accordage par quintes et quartes pures aboutit au 13e son qui est un si#, plus haut que l'octave du son de départ. On entend ensuite : le $si\#_3$ seul ; puis le do_3 suivi du si# qui forme avec lui un intervalle animé de battements ; puis le do_3 et l'octave pure do_3-do_4 ; enfin la succession do_4, $si\#_3$. L'intervalle entre le do_4 et le $si\#_3$ (plus haut) est le comma pythagoricien. [M. C., Cantor]

Son 8.27 – Démonstration du comma enharmonique. L'accordage de trois tierces pures ascendantes : do_3-*mi, mi*-*sol*#, *sol*#-*si*#, aboutit au $si\#_3$, plus bas que l'octave de la note de départ. On entend ensuite : le $si\#_3$, le do_3 et l'accord simultané do_3-$si\#_3$ (battements) ; puis l'octave pure do_3-do_4 ; enfin do_4 suivi de $si\#_3$. L'intervalle $si\#_3$-do_4 est le comma enharmonique. [M. C., Cantor]

Son 8.28 – Pour aboutir à une tierce pure, il faut altérer légèrement les intervalles qui y conduisent. La quinte est réduite et la quarte est agrandie. En toute rigueur, chaque intervalle doit être altéré d'un quart de comma syntonique. Au cours de la succession : *do*, *do-sol*, *sol-ré*, *ré-la*, *la-mi*, on peut entendre l'alternance des battements lents sur les quintes et plus rapides sur les quartes descendantes. La séquence se termine sur la tierce pure *do-mi*. [M. C., Cantor]

Son 8.29 – Comparaisons de gammes chromatiques jouées selon deux tempéraments : tempérament égal et tempérament mésotonique, et selon deux diapasons différents. Toutes les gammes commencent au do_3. Son 8.29a : clavecin, la_3 = 440 Hz ; tempérament égal, puis tempérament mésotonique. Son 8.29b : Cantor, la_3 = 440 Hz ; tempérament égal, puis tempérament mésotonique. Son 8.29c : clavecin, la_3 = 415 Hz ; gamme chromatique en tempérament égal. Son 8.29d : clavecin, la_3 = 415 Hz ; gamme chromatique en tempérament mésotonique. Réalisation M. C. (Sons a et b) et P.-Y. Asselin (Sons c et d). [Archives LAM]

Son 8.30 – Extrait musical joué sur le Cantor : tempérament égal (la_3 = 440 Hz). *Kleines harmonisches Labyrinth*, J.-S. Bach, début de l'Exitus. [M. C.]

Son 8.31 – Extrait musical joué sur le Cantor : tempérament mésotonique (la_3 = 440 Hz). *Kleines harmonisches Labyrinth*, J.-S. Bach, début de l'Exitus. Exemple pédagogique sans lien avec une réalité historique. [M. C.]

Son 8.32 – Deux sons isolés (si_2) de registres différents du même clavecin, joués successivement : sons A et B de la figure 8.43. [M. C.]

Son 8.33 – Les deux sons précédents, A et B, joués simultanément. L'unisson est satisfaisant. [M. C.]

Son 8.34 – Les deux sons A et B avant accordage. Bien que l'écart ne soit que de 3,78 Hz sur le fondamental, les battements sont très agressifs en raison de la richesse spectrale de l'instrument. [M. C.]

Son 8.35 – Une note (la_2 = 220 Hz) jouée sur deux pianos différents. Bases de données : McGill (piano 1); Iowa (piano 2). [M. C.]

Son 8.36 – Exemple permettant d'écouter le compromis à réaliser entre l'inharmonicité du piano et l'agrandissement de l'octave. L'expérience d'écoute porte sur l'appréciation de l'intervalle la_1-la_2 (la durée des sons est assez brève). On entend successivement quatre paires de sons, chacune étant répétée deux fois. Dans la 1re paire les composantes sont harmoniques et l'intervalle d'octave est pur. Les sons des paires 2, 3, 4, possèdent un taux d'inharmonicité de piano standard (ß = 0,0002), telle que la fréquence du 19e partiel correspond à l'harmonique 20 d'un son périodique (voir chapitre 6, § 3.2). L'intervalle d'octave est pur pour la paire 2 ; il est agrandi pour les paires 3 et 4 (voir le tableau 5 ci-dessous). Expérience réalisée avec des sons réels de piano numérisés puis transformés (harmonicité et transposition en fréquence). Taro Mori, thèse, 2000. [Archives LAM]

Tableau 5. Rôle de l'inharmonicité dans l'appréciation d'une octave « juste » au piano (T. Mori, 2000)

Son	Spectre	Accordage	Fréquences	Commentaire
Paire 1	la_1 harmonique la_2 harmonique	Octave pure	110 - 220 Hz	Comparer les paires 1 et 2
Paire 2	la_1 harmonique la_2 inharmonique	Octave pure	110 - 220 Hz	
Paire 3	la_1 harmonique la_2 inharmonique	Octave + 3 cents	110 - 220,39 Hz	Comparer les paires 2 et 3, 2 et 4
Paire 4	la_1 harmonique la_2 inharmonique	Octave + 5 cents	110 - 220,63 Hz	

Commentaires sur l'écoute : la première paire donne l'impression d'un son unique. La sensation de deux sons à l'octave l'un de l'autre n'apparaît qu'avec la deuxième paire. Il faut ensuite comparer les changements de qualité sonore globale que produit la hausse de la note supérieure dans les paires 2, 3 et 4.

Son 8.37 – Séquence d'accordage d'une quinte pure fa_2-do_3 au clavecin. L'accordeur répète la quinte plusieurs fois tout en agissant sur la cheville d'accord de la note supérieure pour ajuster l'intervalle. À la dernière occurrence (14) l'accordeur descend la note supérieure et cale la quinte pure pendant l'extinction du son. L'intervalle est jugé bon 1,4 s après le jeu. Marandas, E., Mostis, K., Gibiat, V., 1998, Actes du colloque *Acoustique et instruments anciens*, CD plage 29. [SFA, Cité de la musique]

Son 8.38 – Extrait du son 8.37 : accordage de la dernière quinte. [*Ibid.*]

CHAPITRE 9

VOIX ET PERCEPTION

1. La voix humaine : un instrument très particulier

Tout au long de cet ouvrage, nous avons présenté des exemples empruntés à la voix parlée ou chantée. La voix est un instrument très particulier que chacun de nous possède et qui, du fait de son importance au sein des relations humaines, participe intimement au développement de la perception sonore.

Très tôt, l'enfant apprend à différencier les voix de ceux qui l'entourent, à capter la musique de la voix (rythme et intonation) qui transmet les sentiments et les émotions, et peu à peu à décoder les formes spectrotemporelles de la parole tout en accordant sa propre voix et son écoute dès qu'il peut en contrôler les modulations pour chanter. Au cours de ces différents processus, la zone spectrale dans laquelle évoluent les formants vocaux[1] fait l'objet d'une attention particulière. La voix humaine tient aussi une place centrale dans la plupart des musiques. Sa tessiture a été et demeure le modèle premier de l'organisation mélodique des instruments et de l'écriture musicale.

On trouvera de nombreux ouvrages qui traitent de l'organe vocal sous des angles très différents. Certains s'adressent aux médecins phoniatres et aux rééducateurs (Le Huche, 1984 ; Cornut, 2009), d'autres aux chanteurs (Ormezzano, 2000), d'autres encore à des chercheurs en acoustique (Fant, 1970 ; Sundberg, 1987 ; Titze, 1994). Parmi les nombreuses méthodes de chant certaines (W. Vennard, 1967 et R. Miller, 1990) intègrent à des degrés divers les données issues de la recherche en acoustique et en physiologie. Toutefois, il nous a paru nécessaire de rassembler quelques données sur le fonctionnement de cet *instrument* sonore exceptionnel, afin de mettre en évidence les rapports singuliers qui se tissent entre les potentialités acoustiques du système phonatoire humain, diversement exploitées selon les cultures, et la richesse des modalités perceptives des sons vocaux.

1.1. Une source acoustique polymorphe

Nous avons dit à plusieurs reprises que les sons de la voix humaine étaient aisément reconnaissables par les auditeurs. À quelles caractéristiques acoustiques faut-il attribuer cette singularité qui les distingue des autres sons que nous entendons ?

Les sons d'origine mécanique relèvent des cinq catégories de base schématisées dans la figure 9.1. On distingue deux modes d'excitation : impulsions et entretien ; et trois types de contenu spectral : large bande plus ou moins « colorée » (A, C et

1. La « zone d'écoute dominante », équivalente à la « zone d'existence » de Ritsma (voir chapitre 3, § 2.9).

D), fréquences discrètes inharmoniques (B) ; fréquences harmoniques (E). La forme temporospectrale caractéristique d'un son quelconque résulte de l'évolution temporelle des variations de l'énergie selon un schéma spectral donné[2].

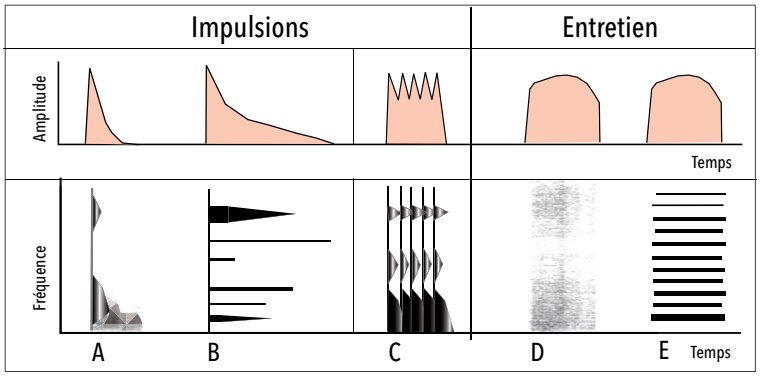


Figure 9.1 Typologie acoustique élémentaire. Partie supérieure : enveloppe du signal temporel ; partie inférieure : représentation schématique du contenu spectro-temporel caractéristique des différents types de productions sonores.

Son 9.1 (5")

Cinq sons de l'environnement

Son 9.2 (6")

Six sons vocaux

À titre d'exemple, le Son **9.1** fait entendre cinq sons extraits de notre environnement, illustrant cette typologie. L'analyse spectrographique en est donnée à la partie supérieure de la figure 9.2. On reconnaît successivement : le débouchage d'une bouteille (A) ; le grincement d'une chaise (C) ; un bruit de frottement (D) ; un klaxon de voiture (E1) ; un glissando d'instrument de musique électronique (E2). Ces cinq sons renvoient à des sources acoustiques de catégories cognitives différentes.

La partie inférieure de la figure 9.2 montre une suite de sons relevant de la même typologie spectrotemporelle, mais, à l'écoute du Son **9.2**, il est manifeste qu'ils relèvent tous d'une seule catégorie cognitive : la voix humaine.

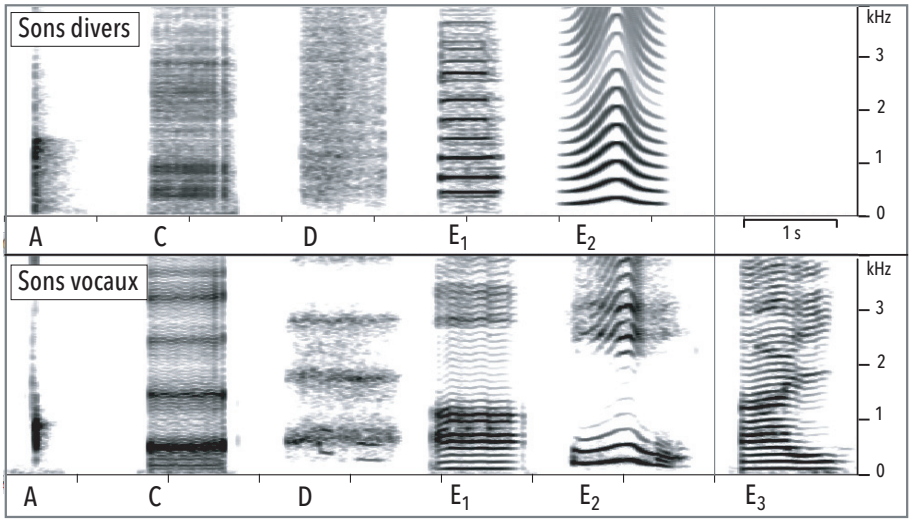


Figure 9.2 Analyse acoustique de quelques sons représentatifs de la typologie exposée sur la figure 9.1. En haut, sons divers (environnement) ; en bas, sons de voix humaine.

2. Voir chapitre 4, § 3.4 ; chapitre 7, § 2.1.

Contrairement aux instruments de musique, dont nous avons vu que chacun d'eux était caractérisé par une forme temporospectrale particulière, résultant de la combinaison d'un type de production sonore avec les modes propres d'une structure vibrante définie, l'instrument vocal humain est une source polymorphe, un organisme vivant dont toutes les parties sont ajustables et peuvent se transformer rapidement sous le contrôle de son hôte, passant de la production d'un bruit de large bande au son périodique, des impulsions au son continu. Or, malgré cette variabilité sonore, la voix humaine est fortement reconnaissable.

Observons les analyses de la figure 9.2. On peut remarquer que tous les sons vocaux ont en commun des zones formantiques (voir Glossaire), véritables « marques de fabrique » du son vocal, dont nous verrons qu'elles sont dues aux résonances de cavités internes. Le son E3, dans lequel les formants varient, indépendamment de la fréquence fondamentale, est tout à fait spécifique de l'instrument vocal. Ces zones spectrales de résonance que nous allons étudier en détail, sont, pour un auditeur, porteuses d'informations différentes selon que celui-ci écoute :

- en mode identitaire (de qui est cette voix ?) ;
- en mode phonétique (sons d'une langue) ou sémantique (écoute de la parole) ;
- en mode qualitatif musical, lequel peut prendre des aspects fort différents, depuis l'appréciation des dimensions expressives d'une voix donnée à celle d'une mélodie formantique comme dans la technique du chant diphonique.

En bref, les sons de la voix humaine ont en commun des indices spectraux spécifiques d'une catégorie acoustique, mais les interprétations auxquelles ils donnent lieu, c'est-à-dire le sens que nous leur attribuons et les divers modes de qualification que nous mettons en œuvre à leur écoute, relèvent de catégories cognitives. Celles-ci dépendent au premier chef des situations dans lesquelles se manifeste la voix : cris, chant, parole. Propres à une culture donnée, elles dépendent aussi, et à des degrés divers, de l'histoire individuelle de chaque auditeur.

Ces divers modes d'appréhension du signal vocal, et plus particulièrement des résonances formantiques, nous poseront à nouveau le problème crucial de l'interprétation des analyses acoustiques que l'on peut très précisément caractériser, alors que l'émetteur – ici le système phonatoire humain – est variable et peu reproductible. Une fois de plus, nous verrons qu'il s'agit moins de cumuler les résultats de mesure sur des paramètres acoustiques indépendants que d'appréhender les rapports de grandeurs qui sont significatives pour l'auditeur et qui donc peuvent rendre compte du traitement holistique des formes sonores spectrotemporelles, de façon analogue à la démarche que nous avons adoptée dans le chapitre qui traite de la perception (4).

1.2. Données élémentaires sur l'instrument vocal

De la structure extrêmement complexe du système phonatoire humain – cartilages, muscles, muqueuses – nous ne retiendrons pour cette présentation que les parties qui jouent un rôle décisif dans la fonction sonore (voir figures 9.3 et 9.4 ; voir aussi Henrich-Bernardoni, 2014).

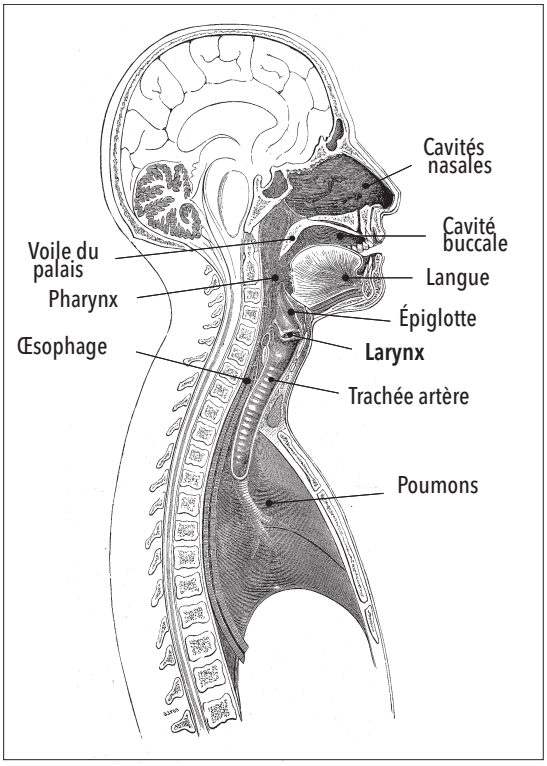


Figure 9.3 Coupe sagittale montrant la position des différentes parties de l'organe phonatoire dans le buste et la tête.

Source : Behnke, E., 1890, planche hors texte.

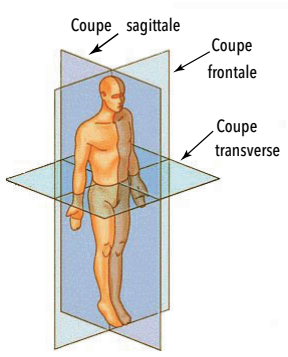


Figure 9.4 Les trois plans de coupe référencés en anatomie.

Source : fr.wikipedia.org/wiki/Anatomie_humaine.

Rappelons tout d'abord que la voix humaine est un instrument à vent.

Certains auteurs (Damsté, 1966 ; Leipp, 1967) ont établi un parallèle entre le fonctionnement acoustique de la voix et celui d'un instrument à embouchure comme le cor. En effet :

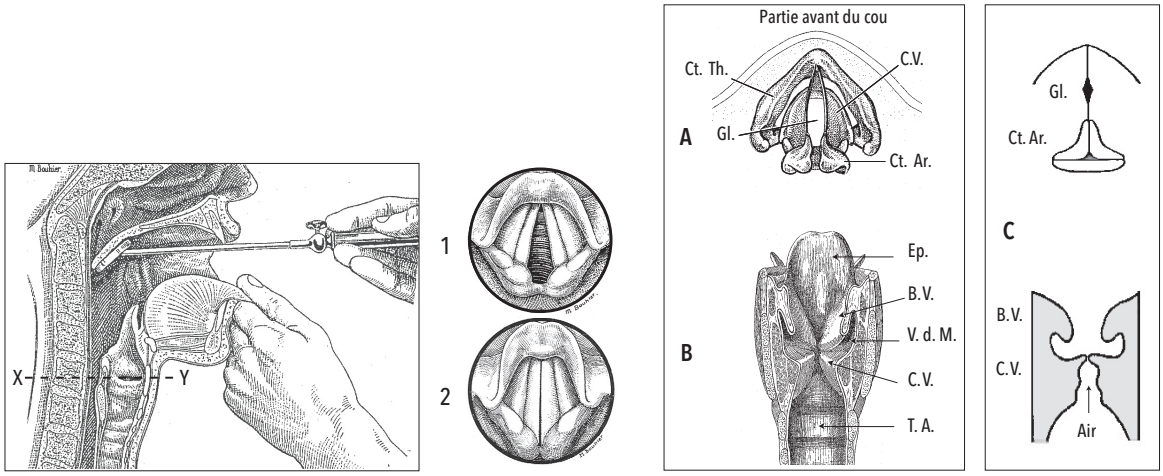
- l'énergie est fournie par l'air pulmonaire expiré sous pression ;
- la source acoustique est due aux vibrations d'un système analogue aux lèvres buccales : les cordes vocales[3] situées dans le larynx ;
- le contenu spectral des ondes sonores initiales est transformé par les résonances des cavités qu'elles traversent (pharynx, cavités buccales et nasales), tout en interagissant avec elles ;
- le rayonnement du son à l'extérieur se produit à la sortie d'un orifice principal, la bouche[4].

1.2.1. Le larynx

Le larynx constitue la partie terminale de la trachée artère. Sa fonction première est de protéger celle-ci des aliments liquides ou solides qui pourraient y pénétrer. Lors de la déglutition, le larynx se ferme et il est recouvert par l'épiglotte (voir figure 9.6-B, Ep.).

La figure 9.5 représente, en coupe sagittale, l'observation du larynx à l'aide du miroir de Garcia ou de tout autre dispositif placé dans l'arrière-gorge (caméra endoscopique, fibre optique).

3. Certains auteurs proposent une nouvelle terminologie : plis vocaux (*vocal folds* en anglais) ou lèvres vocales. L'essentiel est de bien prendre conscience qu'il ne s'agit pas de cordes au sens acoustique, mais de muscles qui peuvent se contracter et être étirés.
4. Les narines n'interviennent qu'épisodiquement et leur contribution énergétique au rayonnement est minime.



De gauche à droite.

Figure 9.5 Observation de la partie supérieure du larynx à l'aide d'un miroir incliné à 45°. À droite, dessins de la glotte ouverte à l'inspiration (1) ; et de la flotte fermée pour la phonation (2).

Figure 9.6 (A) Coupe transversale du cou au niveau de la glotte. (B) Coupe frontale du larynx montrant le profil interne du conduit laryngé. (C) Représentations schématiques de l'ouverture de la glotte entrouverte (en haut), et du profil laryngé correspondant à un état donné de la phonation (en bas). Voir les figures 9.9 et 9.10 pour les variations pendant une période.

Source fig. 9.5 : Tarneaud, J., 1941, figures 12, 13, 14.

Source fig. 9.6 : A : d'après Habermann, G., 1978, figure 21a. B : d'après Garnault, P., 1895, figure 22.

Pendant l'inspiration et l'expiration, le larynx se présente sous l'aspect de la figure 9.5-1 avec une ouverture de forme triangulaire : la glotte (Gl.).

La fermeture de la glotte (figure 9.5-2), est provoquée par le mouvement de deux petits cartilages situés dans la partie postérieure : les cartilages aryténoïdes (Ct. Ar.) qui rapprochent énergiquement les cordes vocales.

La figure 9.6-A représente une coupe transversale du cou passant par le plan des cordes vocales (coupe au niveau X-Y de la figure 9.5), et montrant le cartilage thyroïde (Ct. Th.) et les deux aryténoïdes (Ct. Ar).

On voit que les deux cordes vocales (C.V.) ont un point d'attache commun à l'avant, sur la face interne du cartilage thyroïde. Chez certains sujets, ce cartilage est saillant sur la partie antérieure du cou : c'est la « pomme d'Adam ». Les mouvements ascendants et descendants de la pomme d'Adam, visibles pendant la déglutition, témoignent de la mobilité du larynx qui est un organe souplement suspendu.

La figure 9.6-B montre la forme du profil interne du conduit d'écoulement de l'air. On observe deux rétrécissements séparés par une cavité évasée (le ventricule de Morgagni, V. d. M.). Le rétrécissement inférieur est formé par les cordes vocales, organes actifs de la production sonore ; le rétrécissement supérieur, dû aux bandes ventriculaires[5] (B.V.), n'entre en action que pour des techniques vocales spécifiques, lors de la production de sons de très basse fréquence (technique du *period-doubling*, voir Bailly, 2008).

La figure 9.6-C représente une schématisation des vues 9.6-A et B.

5. Les bandes ventriculaires portent quelquefois le nom de fausses cordes vocales.

1.2.2. La production de sons : de l'écoulement bruité au son de fréquence définie

Les mouvements respiratoires normaux sont silencieux. Un flux d'air bruyant est signe de rétrécissements dans le conduit. La voix chuchotée, par exemple, est une expiration volontairement sonorisée par rapprochement partiel des cordes vocales. Pour que se produise un son de fréquence définie, il faut que les cordes vocales s'accolent complètement afin que s'amorce un cycle périodique d'obturations et d'ouvertures qui module le débit aérien[6], comme décrit dans l'encadré de la page suivante. La rapidité, la régularité et la forme des mouvements glottiques vont déterminer les caractéristiques acoustiques de la vibration, c'est-à-dire la fréquence, la périodicité et le contenu harmonique de la source sonore vocale.

Son 9.3 (4")

Transition voix chuchotée-voix normale sur une voyelle tenue

Dans l'exemple sonore suivant (Son **9.3**), la voix est tout d'abord chuchotée (voyelle « e »), puis les cordes vocales se rapprochent et restent accolées pendant environ 1,5 s, enfin la voix redevient chuchotée. Le son voisé s'établit de façon assez abrupte et termine de même. Il faut beaucoup d'entraînement pour amorcer un son chanté de façon très progressive.

Cette voix comporte du souffle et, sur l'analyse spectrale de la figure 9.7, on remarque que les zones de résonance de la voyelle, dans lesquelles sont concentrées les bandes de bruit, persistent pendant l'émission du son périodique. Certaines restent stables, mais la bande la plus grave baisse notablement lorsque la glotte se ferme pour vibrer. À l'exception du fondamental, les harmoniques les plus intenses (rouges sur la figure) sont les plus proches des bandes de résonance.

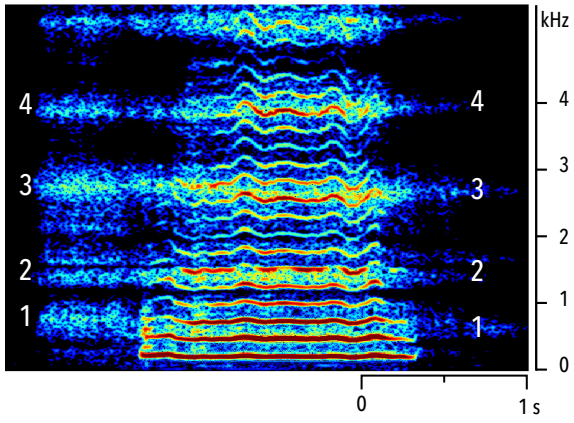


Figure 9.7 Exemple d'une émission vocale passant d'une voix chuchotée à une voix de hauteur définie, sur la même voyelle « e ». Voix féminine.

La qualité finale du son que nous entendons à la sortie de la bouche dépend fortement des transformations que l'onde aérienne initiale, formée à la sortie de la glotte, aura subies au cours de la traversée des cavités qui sont sur son trajet (pharynx, bouche, cavité nasale). Nous touchons là une des difficultés majeures de l'étude acoustique de la voix, car il n'est pas possible d'étudier séparément les caractéristiques de la source de production et celles du corps sonore, comme on peut le faire avec la trompette (par exemple le système lèvres + embouchure, d'une part, et le tuyau d'autre part. Pour donner une idée de l'élaboration acoustique du son vocal, nous prendrons l'exemple de la glottographie.

6. Nous laisserons de côté le ronflement, qui entre dans cette catégorie, mais dont la production (voile du palais, langue) sort de notre propos.

Observation du larynx en mouvement

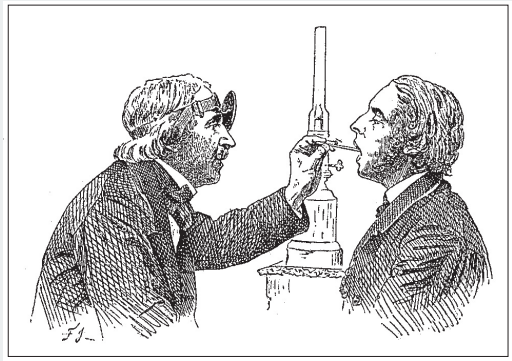


Figure 9.8 Première observation de la glotte avec un miroir.

Garcia, M., 1884, figure 5.

Pour comprendre le fonctionnement du larynx, il faut pouvoir en observer les mouvements. C'est un chanteur, Manuel Garcia Jr., qui fit les premières observations en 1855 (Castellengo, 2005). Il eut l'idée d'utiliser un miroir de dentiste et, en s'éclairant à la lumière du soleil, il décrivit les transformations du comportement des cordes vocales lors des changements de registre. Cependant, le mouvement détaillé de la vibration des cordes vocales échappait à la vue directe, car il est trop rapide. D'immenses progrès ont été faits depuis, grâce à la stroboscopie qui permet de construire une image ralentie de la partie périodique du mouvement, puis à la cinématographie ultrarapide (4000 images/s) qui seule permet d'observer les transitoires (attaque du son, changements de mécanismes). En combinant l'observation directe des cordes vocales avec l'imagerie IRM de l'organe dans le sens de l'écoulement de l'air (profil laryngé), on obtient une représentation intelligible du comportement vibratoire du larynx.

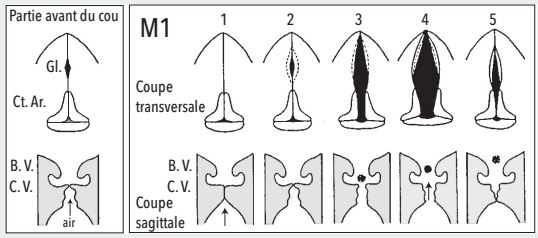


Figure 9.9 Étapes successives d'ouverture et de fermeture du larynx pendant une période de vibration. Mécanisme 1 (M1).

D'après Vennard, W., 1967, figures 32 et 36 ; Cornut, G., 2002, p. 14-15

Avant la phonation, les cordes vocales se joignent pour obturer le larynx (1). La pression de l'air dans la trachée augmente, commence par écarter le bord inférieur des cordes vocales, et une ouverture se forme au milieu (2). Les cordes vocales s'ouvrent sur toute leur longueur (3) et cèdent le passage à une bouffée d'air sous pression. Immédiatement après l'ouverture complète (4), les cordes vocales se rapprochent et se referment en commençant par leur bord inférieur (5). Ce schéma correspond aux mouvements du larynx pour l'émission de sons graves et medium du mécanisme 1 : ceux de la voix dite de poitrine.

1.2.3. De la glotte aux lèvres : la mise en forme spectrale des sons vocaux

L'électroglottographe imaginé par Philippe Fabre en 1957 est un appareil qui délivre un courant de haute fréquence et de faible intensité aux bornes de deux électrodes que l'on applique sur le cou, de part et d'autre du cartilage thyroïde. Le signal recueilli – électroglottogramme ou EGG – est modulé en amplitude par le mouvement d'ouverture et de fermeture des cordes vocales[7] (figure 9.12).

7. Le courant ne passe que lorsque les cordes vocales s'accolent. Le signal EGG reproduit avec précision la période du son.

Mécanismes vibratoires laryngés

Le vibrateur laryngé peut produire des sons couvrant une très large tessiture, de quelques hertz à plus de 2000 dans les cas exceptionnels. Pour y parvenir, le chanteur modifie le souffle et les tensions musculaires, et dispose de plusieurs configurations vibratoires laryngées, chacune étant plus particulièrement adaptée à la production des sons graves, medium, aigus. Trois paramètres principaux régissent la fréquence de vibration d'un système mécanique : la masse, la raideur et la longueur vibrante. Les sons produits dans le mécanisme 1 (M1, figure 9.9) correspondent à la zone 100 Hz-550 Hz (sol_1-$ré_4$), mais avec un effort croissant au fur et à mesure que la fréquence monte. Pour s'adapter à l'aigu, le vibrateur laryngé s'allège et bascule en mécanisme 2 (M2, figure 9.10) : les cordes vocales sont plus fines, plus étirées et la portion vibrante souvent réduite. Une partie importante de la tessiture vocale (fa_2-do_4) est commune aux deux mécanismes (figure 9.11).

Les mécanismes laryngés M1, M2 et M0 (ou voix fry qui génère les sons de l'extrême grave) ont été observés et décrits par les physiologistes depuis le XIXe siècle. Ils sont distincts des "registres" des chanteurs, qui sont des portions de tessiture de qualité vocale homogène, dont certaines sont produites dans le même mécanisme laryngé (voir Roubeau *& al.* 2009).

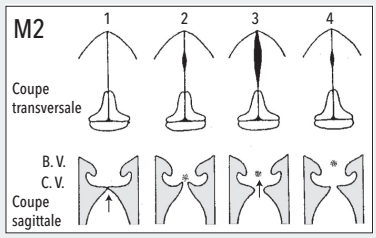


Figure 9.10 Étapes d'une période en mécanisme M2.

D'après Vennard, *op. cit.* et Cornut *op. cit.*

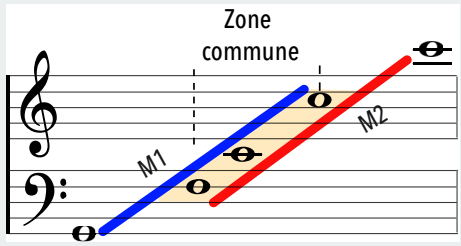


Figure 9.11 Tessiture des mécanismes M1 et M2 avec leur zone commune (les limites figurées sur la portée ne sont qu'indicatives).

Son 9.4 (10")

Signal EGG amplifié

Son 9.5 (10")

Son chanté capté au microphone

Dans l'expérience suivante, un chanteur baryton enregistre une phrase musicale et nous captons simultanément le signal EGG. Écoutons tout d'abord la sonorisation de l'EGG (Son **9.4**). Nous entendons bien la mélodie avec son phrasé et la conduite du vibrato mais le son est étrange : ce n'est pas une voix. Le signal vibratoire recueilli à la glotte, spectralement uniforme, s'apparente plutôt à un signal synthétique. En écoutant le son enregistré devant la bouche (Son **9.5**), il devient manifeste que les résonances bucco-pharyngées sont indispensables à la reconnaissance d'un son vocal. Le contraste entre les deux signaux permet de prendre conscience de l'importance des transformations acoustiques qui se produisent lors de la traversée des cavités de résonance : tout signal vocal en porte l'empreinte, visualisable sur un spectrogramme (figure 9.12).

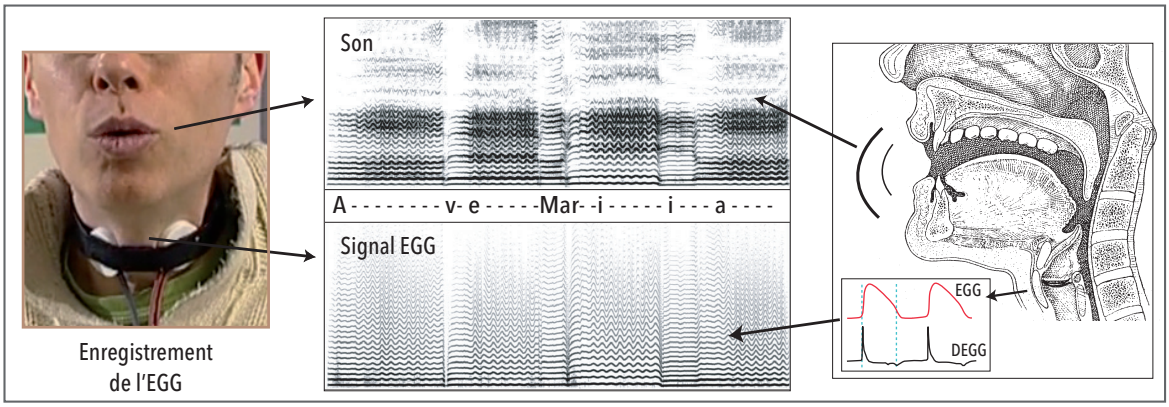


Figure 9.12 Le signal électroglottographique (EGG) capté au niveau du cou, modulé par la succession des ouvertures et fermetures de la glotte, porte les variations de la fréquence fondamentale. Les variations spectrales, en particulier les formants vocaliques, n'apparaissent que sur le son capté à la bouche. Comparez les analyses du son et celle du signal EGG enregistrés simultanément.

Le signal EGG fournit aussi d'importantes indications sur le fonctionnement vibratoire du larynx au cours d'une période, en particulier le quotient Oq, rapport de la durée d'ouverture de la glotte à celle de la période, critère important pour la discrimination des mécanismes vibratoires laryngés (voir Henrich *& al.*, 2003).

1.3. La parole et les articulateurs

Toutes les parties du système phonatoire humain sont d'une extrême mobilité, ce qui offre de grandes possibilités pour modifier temporellement les sons produits par le larynx. Du point de vue acoustique, on distingue :

- des configurations résonantielles correspondant aux voyelles ;
- des événements particuliers correspondant aux consonnes : interruption totale du débit au niveau de la glotte ou des lèvres, interruption brève par la langue, génération de bruits par constriction locale due à la langue ou aux lèvres.

L'étude précise des positions des organes et de la discrimination des sons d'une langue relève de la phonétique[8]. Du point de vue perceptif cependant, la parole résulte d'une succession rapide de mouvements enchaînés dont l'entité de sens est le plus souvent à une échelle temporelle plus grande que celle du phonème. Nous avons déjà exposé au chapitre 4 l'importance de la notion de « forme acoustique spectrotemporelle », en insistant sur le fait qu'une forme peut être anamorphosée dans les dimensions fréquentielles ou temporelles tout en restant reconnaissable, et que son contenu sémantique est indépendant du matériau sonore qui la porte : souffle de la voix chuchotée, harmoniques de la source glottique, d'une guimbarde, ou d'une prothèse vocale électronique. Voir chapitre 4, § 3.6.3, figure 4.14.

1.4. Les cavités de résonance : voyelles et timbre

Les termes « résonance, résonateur » sont définis dans le glossaire. Tout corps possède plusieurs modes vibratoires correspondant chacun à une fréquence de vibration spécifique. Mais, à la différence d'une flûte ou d'un trombone, instruments dans lesquels la fréquence de jeu est celle de l'un des modes vibratoires du

8. Le nombre de caractères principaux de l'alphabet phonétique international (API) est de 118, ce qui permet de couvrir les sons les plus divers. Le français en compte 37 : 16 voyelles dont 4 nasales ; 18 consonnes ; 3 semi-consonnes. Source Wikipedia : *Alphabet phonétique international.*

tuyau, la fréquence fondamentale d'une voix ne dépend que des caractéristiques de la source laryngée, c'est-à-dire de l'interaction entre la pression sous-glottique et les paramètres mécaniques des cordes vocales (masse, raideur, longueur). Les fréquences de résonance des modes propres des cavités – pharynx, bouche, nez – agissent en modifiant la répartition de l'énergie du spectre harmonique laryngé. Ces transformations spectrales, qui font de la voix humaine un instrument de timbre extrêmement complexe à étudier, sont ordinairement codées en langue : ce sont les voyelles.

1.4.1. La réalisation de trois voyelles cardinales : « i », « a », « ou »

En articulant successivement « i », « a », « ou » chacun pourra constater, en introduisant un doigt dans la bouche, que pour le « i », la langue se porte en avant et vers le haut, pour le « a », elle recule et s'aplatit tandis que les lèvres s'ouvrent, enfin, pour le « ou », la langue se place tout à fait en arrière pendant que les lèvres se ferment en s'arrondissant.

LE MAÎTRE DE PHILOSOPHIE.
La voix A se forme en ouvrant fort la bouche, A.
M. JOURDAIN.
A, A. Oui.
LE MAÎTRE DE PHILOSOPHIE.
La voix E se forme en rapprochant la mâchoire d'en bas de celle d'en-haut, A, E.
M. JOURDAIN.
A, E, A, E. Ma foi, oui. Ah! que cela est beau!
LE MAÎTRE DE PHILOSOPHIE
Et la voix I, en rapprochant encore davantage les mâchoires l'une de l'autre, et écartant les deux coins de la bouche vers les oreilles, A, E, I.
M. JOURDAIN.
A, E, I, I, I, I. Cela est vrai. Vive la science!

Figure 9.13 Texte de Molière.
Molière, *Le Bourgeois gentilhomme*, acte II, scène IV. [Bnf]

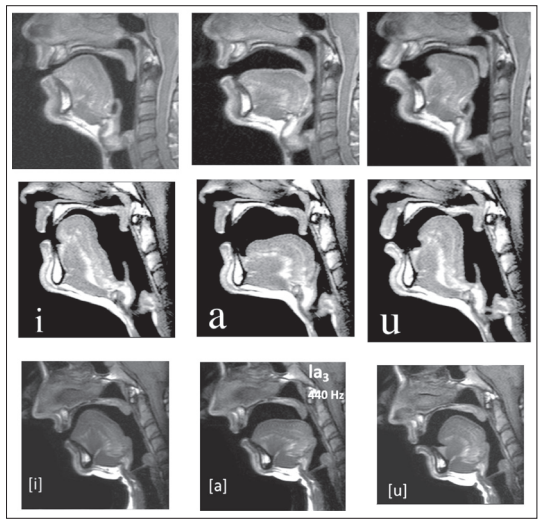


Figure 9.14 Images IRM de la position des organes articulatoires pendant l'émission de trois voyelles. De bas en haut, deux hommes et une femme. Observez la position de la langue et l'ouverture des lèvres.
Crédits : Badin, P., Henrich, N., [Gipsa-lab]

Ces mouvements, connus de longue date (figure 9.13), peuvent maintenant être observés avec une grande netteté grâce à l'imagerie par résonance magnétique IRM (figure 9.14). On voit que le rôle de la langue est complexe. Elle réduit la section du conduit vocal en certains endroits, ce qui a pour effet de délimiter des cavités couplées, plus ou moins bien définies. La cavité postérieure, le pharynx, se situe entre les cordes vocales et le premier rétrécissement du conduit. La seconde correspond à la partie antérieure de la cavité buccale, comprise entre l'orifice des lèvres et le point de rétrécissement précédent (ou arrière).

Que l'on prenne pour modèle du conduit vocal un tuyau de section variable ou une série de résonateurs, les deux paramètres principaux qui déterminent la fréquence de résonance d'une cavité sont : le volume (V) de la cavité et le rapport (S/l) de la section (S) du trou de sortie à la longueur (l) du conduit de raccordement. Ces deux paramètres ont des effets opposés sur la fréquence de résonance. La fréquence diminue quand (V) augmente ou quand le rapport (S/l) diminue[9]. En examinant la réalisation de trois voyelles sur la figure 9.14, on peut voir qu'il existe, malgré les différences anatomiques individuelles entre les trois locuteurs, une bonne convergence des mouvements que nous venons de décrire.

L'observation des images IRM réduites à deux dimensions donne des indications incomplètes mais, dans le cas extrêmement contrasté des trois voyelles choisies, l'interprétation des figures permet d'établir une correspondance entre la réalisation articulatoire et les résonances visibles sur l'analyse acoustique.

1.4.2. La caractérisation acoustique des voyelles : les formants spectraux

Nous avons demandé à un locuteur masculin de varier la fréquence fondamentale de sa voix tout en s'efforçant de maintenir constante chacune des trois voyelles. Dans l'exemple sonore **9.6**, on entend successivement « i », « a », « ou » émis avec un glissando rapide (ascendant-descendant) de la fréquence fondamentale.

Son 9.6 (8")

Les trois glissandos de la figure 9.15 (homme)

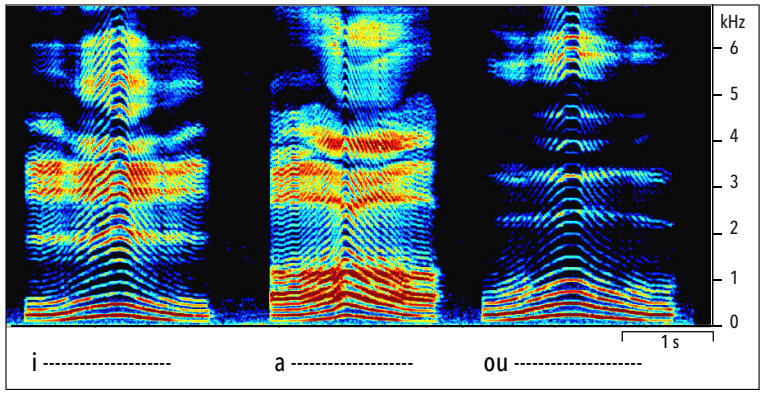


Figure 9.15 Trois voyelles, « i », « a », « ou » chantées chacune avec un glissando ascendant-descendant rapide. Voix masculine.

Sur l'analyse de la figure 9.15 apparaissent de nombreuses zones horizontales correspondant aux résonances qui sont à peu près fixes pendant les variations de l'intonation : on en compte plus de cinq du grave à l'aigu. Toutes contribuent à la caractérisation acoustique de la voix, mais les études de la parole, et en particulier les premiers essais de synthèse, ont montré que les résonances les plus graves avaient un rôle prépondérant dans l'identification des voyelles. Les résonances se traduisent par des zones de renforcement spectral qu'on nomme « formants ». En pratique, ce sont les deux premiers formants qui sont pris en compte pour la caractérisation des voyelles.

9. Un exemple de calcul de la fréquence d'un résonateur est proposé annexe C.

Son 9.7 (5")

Transition continue « i, a, ou » en voix fry : homme

Son 9.8 (5")

Transition continue « i, a, ou » en voix fry : femme

Dans les deux exemples sonores suivants, nous avons demandé à un homme (Son **9.7**) et à une femme (Son **9.8**) de produire la même succession de voyelles enchaînées avec l'émission très grave, quasi rauque, de la voix « fry »[10], qui a pour avantage de faire disparaître les raies harmoniques dont nous verrons qu'elles compliquent le repérage des formants sur les analyses.

La figure 9.16 montre très clairement le trajet spectral du premier formant (F1) qui évolue (pour l'exemple masculin) entre 300 et 800 Hz, et celui du deuxième formant (F2) qui descend de 2000 Hz pour le « i » vers 600 Hz pour le « ou » (/u/ en code phonétique). Les variations spectrales de la voix féminine sont tout à fait similaires, mais transposées vers les fréquences plus aiguës.

Il est intéressant de situer les voyelles dans un plan dont les axes sont les fréquences de F1 et F2 (figure 9.16). Les voyelles « a », « i », « ou » de notre exemple forment deux triangles distincts, l'un pour la voix d'homme, l'autre pour la voix de femme.

En effet, les rapports de fréquence des voyelles entre elles étant équivalents, les caractéristiques formantiques d'une voix d'homme et d'une voix de femme sont similaires à la transposition près. Celle-ci est de l'ordre d'une tierce majeure (soit un rapport de fréquence d'environ 5/4).

Cet écart permet de discriminer statistiquement une voix de femme d'une voix d'homme ayant la même hauteur fondamentale. Ce n'est qu'une indication, car il existe de nombreuses exceptions à cette règle.

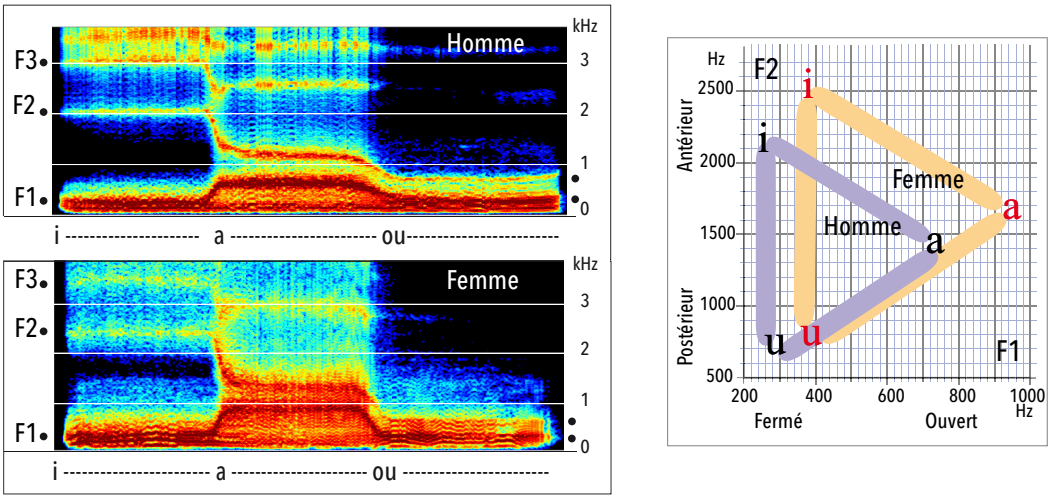


Figure 9.16 Analyses spectrographiques de l'enchaînement des voyelles « i », « a », « ou » émises en mécanisme 0 (voix « fry ») par un homme et par une femme. En vis-à- vis, position des trois voyelles dans le plan des deux premiers formants F1 et F2, pour des voix d'homme et de femme.

10. Émission très relâchée qui intervient fréquemment en fin de phrase, sur l'hésitation « euh... ». Elle est caractérisée comme mécanisme 0.

1.4.3. Les formants de 11 voyelles du français : voix d'hommes, émission en fry

À l'intérieur du triangle formé par les trois voyelles cardinales (i, a, ou) se placent les autres voyelles (figure 9.17), en nombre variable selon les langues et, pour une langue donnée, selon les accents territoriaux (voyelle plus ou moins ouverte ou fermée).

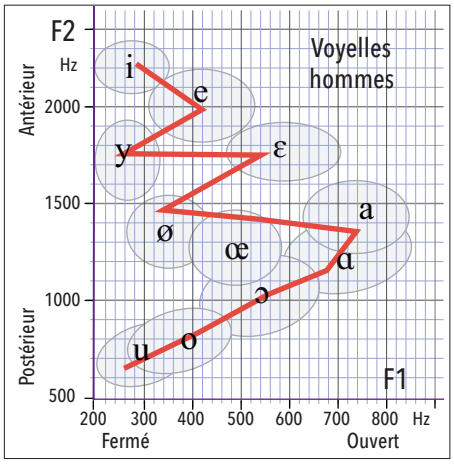


Figure 9.17 Positions relatives des voyelles du français représentées dans le plan F1-F2. Moyennes sur un corpus d'hommes canadiens.

Données expérimentales : Université de Laval, www.phonetique.ulaval.ca.

La figure 9.18 montre l'analyse spectrale des voyelles articulées par un locuteur masculin en voix fry (Son **9.9**), ordonnées selon la valeur de F2 (trait en pointillé rouge) qui décroît de « i », fréquence la plus élevée, jusqu'à « ou », fréquence la plus basse. On note que la fréquence de F2 est différente pour chaque voyelle. Il existe de petites différences entre l'analyse spectrographique et les positions des voyelles du graphique, car celui-ci rend compte de mesures statistiques de voix masculines du Québec. L'analyse spectrographique met en évidence l'opposition entre le groupe des voyelles antérieures (langue en position avant), sur la partie gauche du sonagramme, et les voyelles postérieures (langue vers l'arrière) à la partie droite. Dans le groupe « a », « o », « ou », les formants F1 et F2 évoluent de façon parallèle, ce qui offre d'intéressantes possibilités musicales (voir § 2.3.3).

Son 9.9 (10")

Les voyelles d'un locuteur français masculin enchaînées continument en voix fry (mécanisme M0)

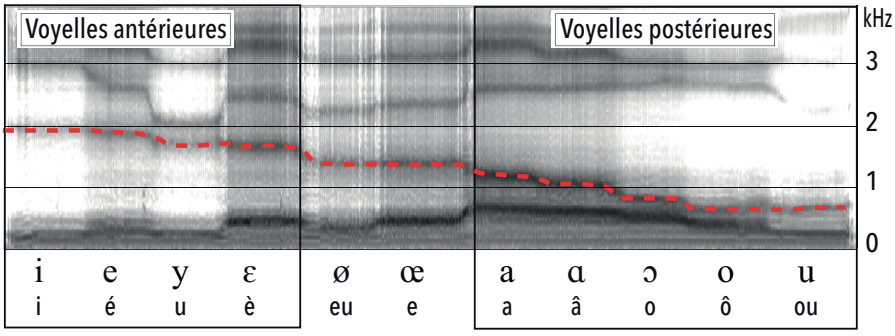


Figure 9.18 Analyse spectrographique de 11 voyelles orales prononcées par un locuteur français, masculin, en voix fry. Le trait pointillé rouge surligne le trajet du formant F2.

2. La voix chantée et les cavités de résonance

Dans ses manifestations les plus diverses, le chant privilégie le son de hauteur définie produit par l'oscillation des cordes vocales. Lorsque le chant comporte des paroles, les parties consonantiques qui perturbent la vibration périodique sont considérablement amoindries, tant en intensité qu'en durée. Certains styles de chant en sont même totalement dépourvus. L'ajustement précis des cavités de résonance en rapport avec la fréquence fondamentale de la voix est une des clés de la technique vocale, car il a un retentissement global en termes de qualité du timbre, de confort vocal et même de justesse. Ce point est d'autant plus critique que les ajustements changent pour chaque valeur de la hauteur fondamentale de la voix, et que leur mise en pratique diffère selon qu'il s'agit de voix graves ou aiguës et, à même hauteur, selon qu'il s'agit de voix masculines ou de voix féminines.

2.1. Interactions d'un son harmonique avec un résonateur

Tout d'abord, rappelons brièvement qu'un résonateur traversé par un son quelconque transforme la répartition de l'amplitude sans apporter d'énergie supplémentaire (voir Liénard, 1977, p. 77, et l'entrée *Résonateur* du Glossaire). Un résonateur agit comme un filtre en réduisant l'amplitude de certaines composantes, au profit de celles qui sont en accord avec ses fréquences de résonance. Cet effet est d'autant plus marqué que le résonateur a un faible *amortissement*, paramètre qui gouverne simultanément la durée de la vibration, quand cesse l'entretien, et la largeur de bande de la résonance.

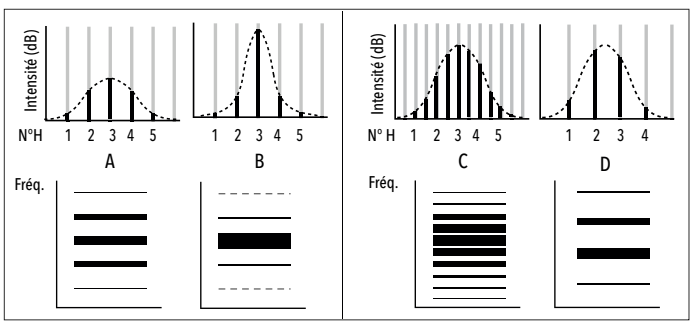


Figure 9.19 schéma montrant l'interaction entre un son harmonique et un résonateur. Modelage du spectre par le résonateur (partie supérieure) et traduction graphique dans le plan sonagraphique (partie inférieure). A et B : variation de l'amortissement du résonateur. C et D : variation de la fréquence fondamentale du son.

Sur la figure 9.19, le résonateur A est plus amorti que B. Ce dernier a une bande passante étroite et donc une grande *sélectivité* : l'harmonique 3 est considérablement renforcé au détriment de ses collatéraux. En ce qui concerne la voix, cette situation exceptionnelle exige un contrôle très fin des articulateurs : c'est le chant diphonique ou harmonique, déjà analysé au chapitre 6. Le plus souvent, les résonateurs buccaux agissent sur plusieurs harmoniques, ce qui ne permet pas de connaître précisément la fréquence centrale de la résonance. Pour un amortissement donné, son estimation est meilleure quand le son est de basse fréquence (C comparé à D). La mesure précise des résonances du conduit vocal pendant le chant nécessite une méthodologie particulière mise en œuvre récemment (voir Henrich, 2011).

2.2. Relations entre la tessiture des voix chantées et les zones spectrales des formants vocaliques

Pour faciliter l'appréhension musicale de cette importante question, nous avons rassemblé sur un même diagramme l'étendue des notes fondamentales des voix chantées et les positions relatives des formants de différentes voyelles, ordonnées selon F1 (figure 9.20).

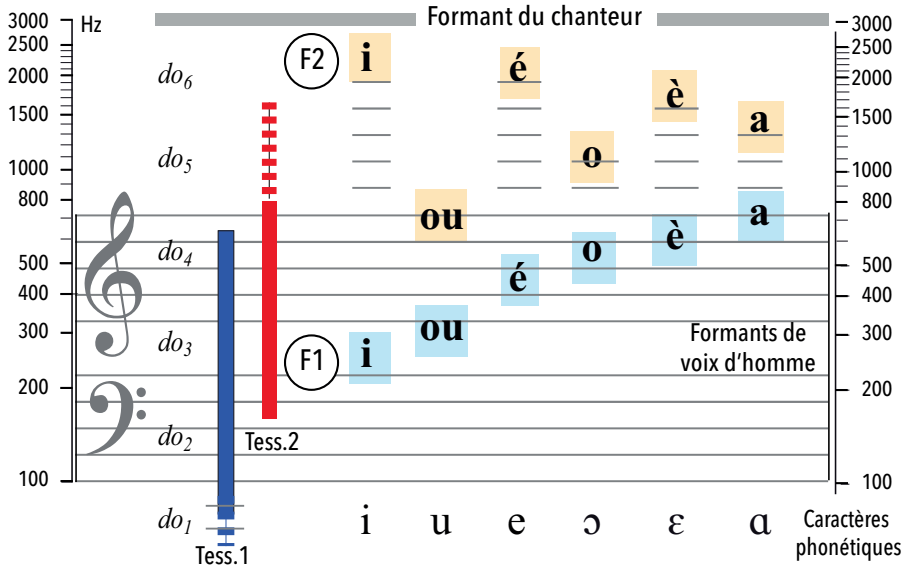


Figure 9.20 Tessiture vocale et zones formantiques des voyelles. De gauche à droite : axe des fréquences selon une échelle logarithmique ; numérotation des notes *do* ; tessiture T1 des voix graves et médiums ; tessiture T2 des voix médiums, aiguës et suraiguës. Position musicale des formants F1 et F2 de six voyelles orales du français parlé pour une voix d'homme. Les positions en fréquence sont indicatives, car la réalisation d'une voyelle admet une certaine tolérance.

L'échelle des fréquences est logarithmique, pour s'ajuster à la portée musicale (voir chapitre 2 § 3.7). La tessiture T1 concerne les voix les plus graves pratiquant le « doublement de période » (zone en pointillé au-dessous du do_1) et l'étendue couverte par le mécanisme M1. La tessiture T2 concerne le mécanisme M2 et sa partie suraiguë (en pointillé) est appelée voix de sifflet.

Comme nous l'avons vu figure 9.11, l'étendue fa_2-do_4 peut être produite en M1 ou en M2. Tant que la fréquence fondamentale de la voix se tient dans les basses fréquences, les formants F1 et F2 affectent principalement l'amplitude des harmoniques supérieurs.

À partir de 250 Hz environ (do_3), la fréquence fondamentale entre dans la zone de fréquence du premier formant, ce qui produit d'une part un changement de qualité du son vocal, d'autre part une modification d'émission ressentie par le chanteur, d'autant que la première résonance (F1) concerne principalement la partie arrière du pharynx, celle qui est proche des cordes vocales.

Son 9.10 (37")

Voir page 472

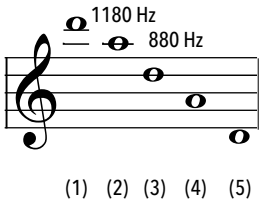


Figure 9.21 Notes sur lesquelles un texte donné est chanté de l'aigu au grave par une soprano.

Pour une note donnée, les effets changent fortement selon la voyelle. Par exemple, en chantant successivement « i-ou-i-ou » sur la note *ré*$_3$, on peut entendre l'harmonique 3 (*la*$_4$) émerger sur le « ou ». Les chanteurs lyriques peuvent ajuster la hauteur d'un formant pour renforcer un harmonique donné – ce que l'on désigne en anglais par *formant-tuning* –, particulièrement dans la zone de 250 à 800 Hz, afin de conserver l'accord F1 sur la fréquence fondamentale. Dans l'extrême aigu de la tessiture, toutes les voyelles finissent par se ressembler et l'intelligibilité devient problématique. L'exemple sonore suivant (Son **9.10**) est éloquent : il vaut mieux connaître d'avance le texte que chante la soprano entre *la*$_4$ et *ré*$_5$. La tessiture de la voix parlée et celle des chansons « à texte », se tient généralement entre 150 et 400 Hz (*ré*$_2$-*sol*$_3$).

2.3. Voyelles et chant lyrique

Son 9.11 (10")

Gamme diatonique de do$_2$ à do$_3$; voyelle « é » ; baryton

En voix chantée, les rapports de fréquence entre la hauteur fondamentale et les fréquences de résonance du conduit vocal se modifient sans cesse, ce qui produit d'importantes variations spectrales affectant les qualités de la voix dans son ensemble : aussi bien la voyelle que l'homogénéité du timbre. Prenons l'exemple d'un chanteur qui travaille une gamme en gardant constante la voyelle « é » : figure 9.22 et Son **9.11**.

2.3.1. Exemple d'une gamme diatonique chantée sur une voyelle tenue (baryton)

Au cours de la montée, la fréquence fondamentale varie alors que les formants de la voyelle « é » restent stables. En plaçant côte à côte les spectres de trois sons, *do*$_2$-*sol*$_2$-*si*$_2$, on peut voir que les changements spectraux des premiers harmoniques sont flagrants. De *do* à *mi*, l'harmonique dominant correspondant à F1 passe de H3 à H2. À partir du *sol,* le chanteur adapte la fréquence du formant F1 pour garder l'accord avec l'harmonique 2. Le formant F2, dont la zone de fréquence est plus large que F1, bouge peu. Le formant F3, situé aux alentours de 3 kHz, est le formant du chanteur déjà vu au chapitre 5 § 2 et au chapitre 7 § 4.2.4.

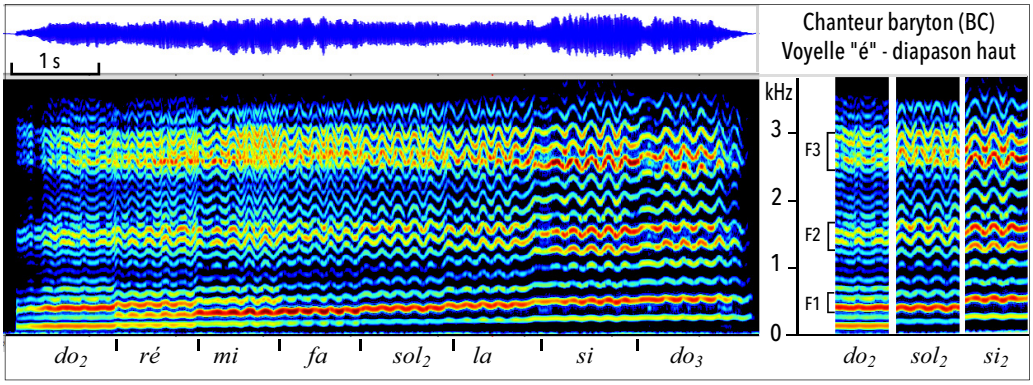


Figure 9.22 Gamme diatonique chantée sur la voyelle « é », de *do*$_2$ à *do*$_3$. Notez les variations spectrales qui se produisent du grave à l'aigu. Partie droite de la figure : le spectre de trois sons extraits de la gamme montre l'importance de ces variations. Remarque : l'échelle d'intensité est représentée par la largeur des traits et par la couleur : bleu (faible) jaune (moyen) et rouge (fort).

À l'écoute, la voyelle ne change pas et la voix paraît homogène. Toutefois, en concentrant fortement son attention sur le son, et en oubliant la voyelle, on peut entendre des nuances de sonorités : un son assez corsé dans le grave, puis plus retenu, voire feutré de *fa* à *la,* et un changement net à partir du *si*[11]. Une fois de plus, nous constatons que la permanence perceptive de stabilité de la voyelle et celle de la continuité de timbre du chanteur sont des constructions cognitives acquises par la pratique de l'écoute.

2.3.2. L'ajustement des voyelles pour modifier finement les qualités du son vocal

À l'inverse de l'exemple précédent, il est particulièrement difficile pour les chanteurs de nous donner la sensation de changer de qualité sonore sans altérer la *justesse* de la voyelle. Cette question fait l'objet de nombreux exercices et tient une place importante dans les méthodes de chant, car sa maîtrise est difficile. Nous donnerons seulement un exemple pour en faire comprendre l'importance aux non-chanteurs.

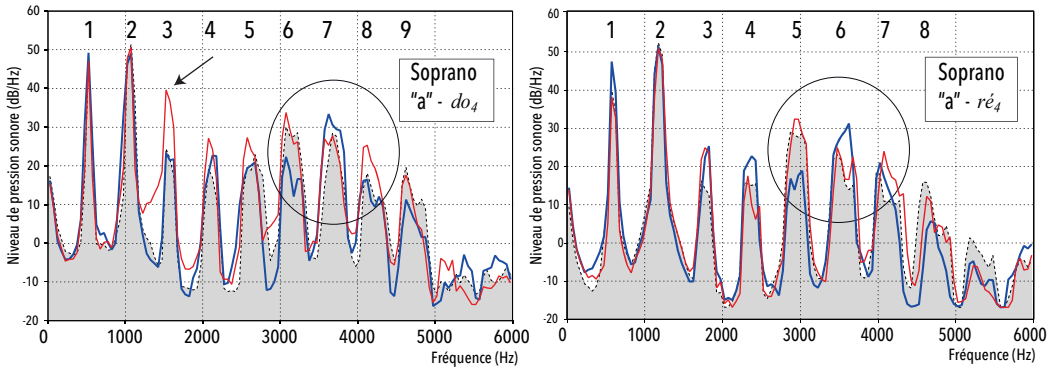


Figure 9.23 Deux exemples de changement de sonorité vocale sur une note, en gardant la même voyelle (soprano). Spectres moyennés des premiers harmoniques. Zone ombrée : son normal ; courbe bleue : son couvert ; courbe rouge : son clair.

Dans les Sons **9.12** et **9.13**, la chanteuse propose l'écoute d'un son chanté sur la voyelle « a » avec deux variations de qualité sonore. On entend le son dit normal, ensuite les sons qualifiés de *couverts* et *clairs*, termes en usage en pédagogie du chant. Dans le premier essai (note do_4), un changement de la voyelle est perceptible sur le son clair.

Son 9.12 (18")

Do_4 : *son normal, couvert, clair*

L'analyse (figure 9.23, flèche) montre une augmentation de 18 dB sur l'harmonique 3 (F2). Dans le deuxième essai (note $ré_4$) la voyelle « a » reste perceptivement stable. Les changements de qualité vocale correspondent à de faibles variations de l'amplitude des premiers harmoniques[12], et surtout à des variations dans la zone de 3 à 5 kHz (cercles sur la figure). La maîtrise consiste donc à modifier l'énergie dans cette région sans perturber la zone des formants F1 et F2 de la voyelle, et ceci pour chaque hauteur de son et pour chaque voyelle.

Son 9.13 (19")

$Ré_4$: *son normal, couvert, clair*

11. On remarque en particulier les harmoniques H6 et H10 qui forment, avec H2, une sonorité de Sesquialtera.
12. À rapprocher de l'analyse des sons « détimbrés » (chapitre 7, § 4.2.4) qui montre l'importance de la balance d'intensité H2-H1, ici très stable.

2.3.3. Le formant-tuning et l'accordage de F1 et F2 en rapport d'octave

Son 9.14 (3'35")

Chant de Pasi but but *pour la pousse du millet (Taïwan).*

Avant d'aborder les mélodies harmoniques proprement dites, nous souhaitons présenter des chants de style intermédiaire qui, tout en préservant la mélodie principale du son glottique s'attachent à produire une qualité vocale particulière en accentuant la résonance de certains harmoniques. Nous avons choisi le *Pasi but but* des Bunun, ethnie aborigène de l'île de Taïwan[13]. Ce chant est remarquable à beaucoup d'égards et il est préférable, avant tout commentaire, de prendre le temps de l'écouter (Son **9.14** en ligne à l'adresse Internet : archives.crem-cnrs.fr ou chant similaire dans *Les Voix du monde*, CDIII-6).

Le *Pasi but but* fait partie des rites pratiqués pour favoriser la pousse du millet, à une époque particulière du calendrier lunaire. C'est un chant cyclique qui commence invariablement dans le grave de la tessiture et monte progressivement et insensiblement, par le truchement d'intervalles glissants, jusqu'au moment ultime où, au dernier cycle, il se termine sur un intervalle de quinte stable, symbolisant la fusion des voix et la pleine lune parfaite.

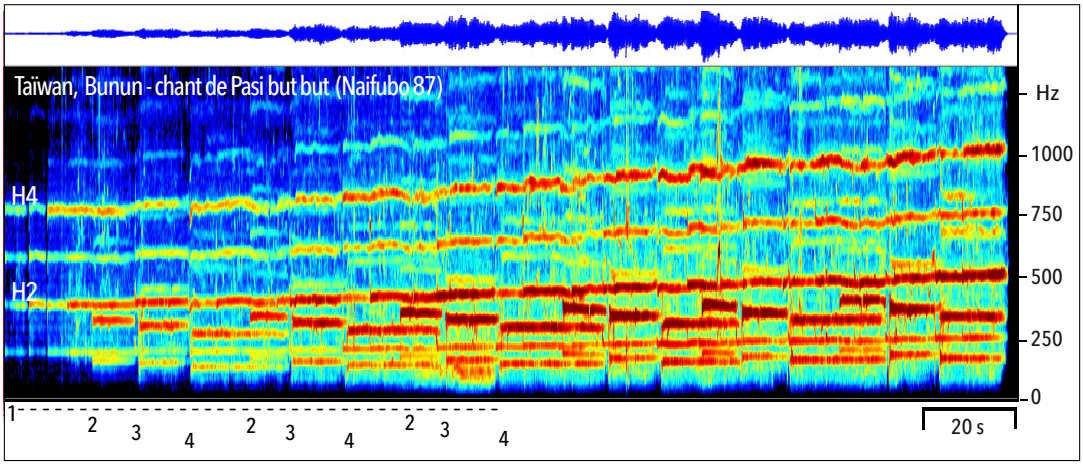


Figure 9.24 Sonagramme du chant de *Pasi but but*. Les numéros 2, 3, 4 indiquent l'entrée successive des voix inférieures à la voix 1. L'ensemble des voix monte graduellement d'une quarte sur la durée du chant.

La voix supérieure (voix 1) est la plus importante. Elle est chantée par trois hommes qui se relaient pour que le son soit continu sur toute la durée du chant, en maintenant la voyelle « o » pendant la montée. Dans notre exemple le fondamental monte de 192 à 255 Hz, soit une quarte (sol_2 à do_3 pour un diapason bas de 1/4 de ton). La « couleur » de la voyelle passe du « o » fermé au « o » ouvert. Pendant que la voix principale monte, les autres voix entonnent successivement la tierce (voix 2), la quarte (voix 3) et la quinte (voix 4) inférieures en utilisant les voyelles « é » et « i » (figure 9.25).

13. Wu Rung Shun (1996), *Tradition et transformation – Le* pasi but but*, un chant polyphonique des Bunun de Taïwan*, Thèse de l'université Paris X, Nanterre, Paris.

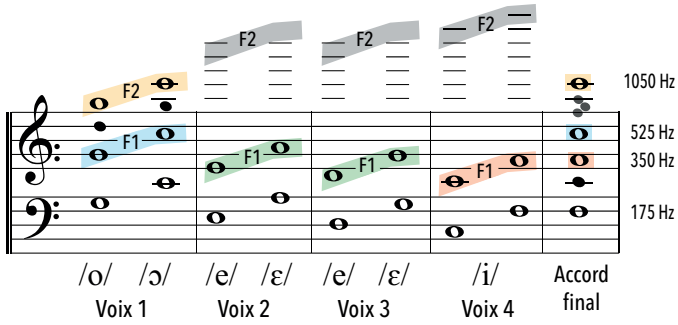


Figure 9.25 La position des formants des voyelles utilisées montre le renforcement de H2 pour toutes les voix, et spécifiquement de H2 et H4 pour la voix 1. Les voix fusionnent sur l'accord terminal de quinte avec renforcement des harmoniques 2, 3, 6 de la fondamentale fa_2.

L'analyse spectrale du chant (figure 9.24) témoigne d'un art consommé du *formant-tuning*. Les harmoniques 2 et 4 de la voix supérieure sont renforcés tout au long du chant, car la voyelle « o » est située sur une droite remarquable (pointillé rouge de la figure 9.26) pour laquelle les fréquences des formants F1 et F2 peuvent être ajustés à l'octave l'un de l'autre (voir les deux zones colorées de la figure 9.26).

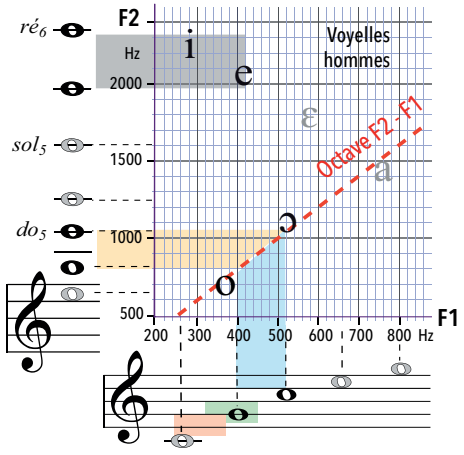


Figure 9.26 Diagramme des voyelles avec correspondance musicale des formants F1 et F2. Le continuum « ô-o-a » est une zone particulière dans laquelle il est possible de se déplacer en tessiture tout en maintenant les deux formants à intervalle d'octave. C'est le cas de la voix 1 du *Pasi but but* et de quelques chants harmoniques : voir § 3.4.

Par ailleurs, l'harmonique 2 des voix 2, 3, 4, également renforcé par F1, reste toujours au dessous de la voix 1 et le deuxième formant de leurs voyelles (« é » et « i ») étant très aigu (vers 2 000 Hz), ces voix n'interfèrent pas avec la zone spectrale de la voix supérieure, qui reste ainsi toujours dégagée. L'accord final du 6e cycle (figure 9.25 à droite) se termine sur la quinte fa_2-do_2 renforcée à l'octave par fa_3 (H2 de la voix 4) et par do_4 et do_5 (F1 et F2 de la voix 1). Cet art sonore des voyelles, pratiqué de diverses façons dans nombre de musiques traditionnelles polyphoniques, est rarement noté sur les transcriptions musicales.

2.1.4. Les jeux de croisement entre mélodies spectrales et mélodies intonatives

Son 9.15 (20")

Extrait de chant style gidayu, normal puis ralenti deux fois

Par les deux possibilités qu'elle offre de faire varier simultanément la hauteur fondamentale et la hauteur spectrale, la voix peut se prêter à des jeux mélodiques ambigus qui sollicitent de façon complexe notre écoute (voir chapitre 6 § 2.7). Dans la musique japonaise, les croisements entre une intonation ascendante et une variation formantique descendante (par exemple de « i » vers « ou ») sont fréquents, et même délibérément recherchés comme dans cet exemple vocal extrait d'une déclamation de style gidayu (Son **9.15** et figure 9.27). La réalisation en est si complexe que nous proposons aussi une écoute au ralenti.

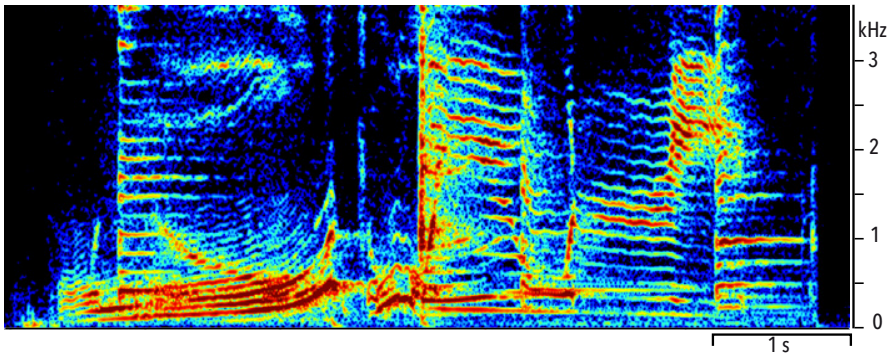


Figure 9.27 Analyse d'un extrait de chant de style gidayu riche en croisements contraires entre la fréquence fondamentale et les mouvements spectraux des formants vocaliques. (Voix et luth, Japon)

3. Voyelles et chants harmoniques : formants F1 et F2

3.1. Les mélodies harmoniques du chant diphonique

Les mélodies harmoniques, qui ont déjà été présentées au chapitre 6, relèvent de techniques tout à fait opposées aux exemples précédents. La source glottique porteuse du spectre harmonique reste stable et le matériau vocalique n'est plus au service d'un texte mais de la seule mélodie. Pour donner à entendre une note, le chanteur augmente la sélectivité de la résonance et en ajuste précisément la fréquence. Partant de l'articulation d'une voyelle, il modifie de façon substantielle la position de la langue et développe ainsi une technique spécifique, différente selon qu'il sélectionne l'un ou l'autre des deux formants principaux. Le choix du formant entraîne aussi celui de la note fondamentale, car c'est le rapport entre cette fréquence fondamentale et celles des harmoniques situés dans la zone de variation du formant qui conditionne l'échelle mélodique disponible (chapitre 8, § 1.3.1).

3.2. Les rapports de fréquence entre formants et fondamentale laryngée

Chant de F2 et chant de F1. L'exemple suivant (Son **9.16**), réalisé par Trân Quang Hai, musicien qui travaille depuis longtemps les techniques du chant harmonique, est particulièrement intéressant, car il permet d'entendre, sur la même note fondamentale do_2 = 133 Hz, la réalisation de deux glissandos harmoniques, l'un par variation du formant F2 et l'autre du formant F1 (voir *Tranquanghai.info).*

Son 9.16 (15")

Chant harmonique utilisant F2 puis F1 (Trân Quang Hai)

La tessiture couverte par les formants d'un homme s'étend environ de mi_3 à la_4 pour F1, et de sol_4 à do_6 pour F2).

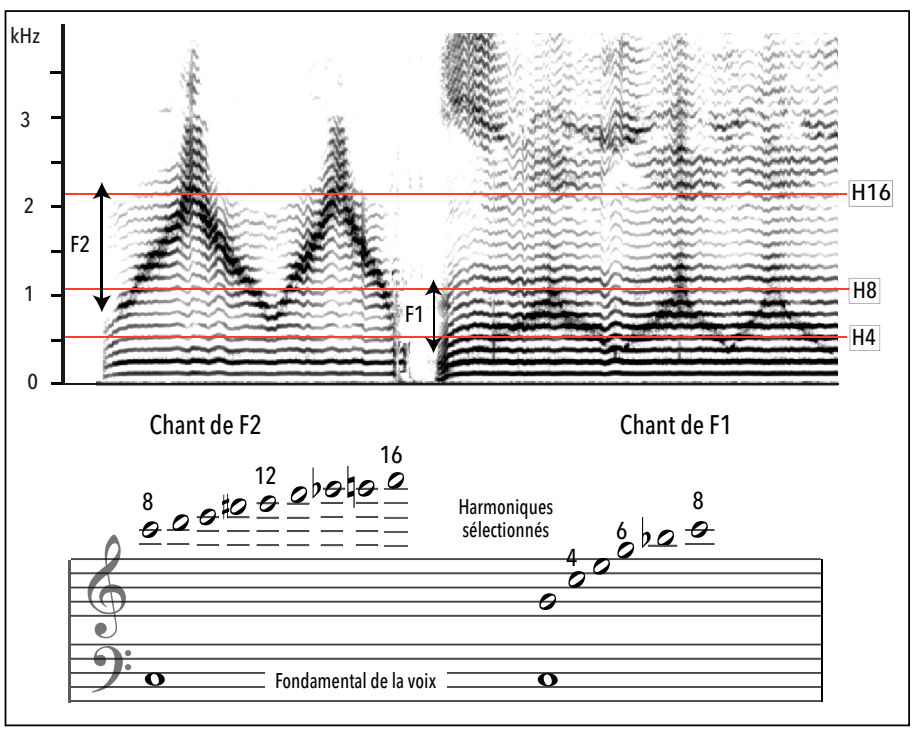


Figure 9.28 Exemple des deux techniques de chant harmonique, utilisant soit F2 et la tranche harmonique 8-16, soit F1 et la tranche harmonique 3-8, sur la même note fondamentale laryngée (do_2 = 133 Hz). Analyse sonagraphique et notation musicale des harmoniques sélectionnés.

Dans cet exemple, le chanteur explore l'étendue de chaque formant en effectuant un glissando résonantiel et l'on entend (Son **9.16**) que l'effet perceptif est musicalement très différent, tant par la qualité vocale et la tessiture que par les possibilités mélodiques. Pour le chant de F2, plus aigu, plus intense à l'oreille, le musicien dispose de 8 harmoniques à l'intérieur de l'octave do_5-do_6 avec un fondamental à do_2 (figure 9.28). C'est le type de chant le plus répandu, connu en Mongolie sous le nom de *Xhoomij* (voir chapitre 8, Son **8.3**, figures 8.6 et 8.7). Les joueurs de guimbarde utilisent aussi les résonances du deuxième formant. Il faut réécouter l'exemple du chapitre 1 (Son **1.20**) et celui du chapitre 8 (Son **8.5**).

Sur la même fondamentale do_2 = 133 Hz, le chant de F1 est comparativement plus sourd. Il évolue entre les harmoniques 4 et 8, ce qui en réduit les possibilités mélodiques. Pour retrouver un nombre équivalent d'harmoniques dans le chant de F1, les chanteurs émettent une fréquence laryngée à l'octave inférieure en adoptant la technique du *period-doubling*.

3.3. Chant de F1 et technique vocale du period-doubling

Ce type de phonation met en jeu les bandes ventriculaires ce qui a pour effet d'augmenter l'inertie du système vibrant et d'abaisser la fréquence vibratoire[14]. En effectuant un zoom temporel sur le début du son de l'exemple de chant kargiraa présenté plus loin, on peut voir nettement le dédoublement de fréquence (voir figure 9.29 et Son **9.19**).

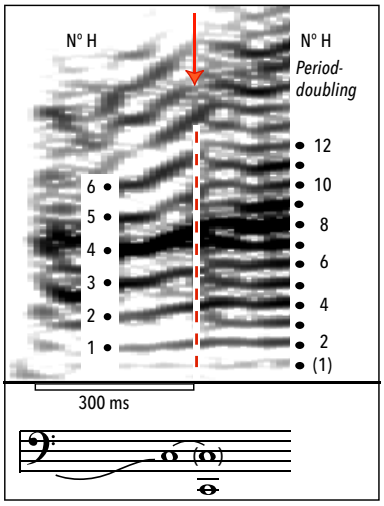


Figure 9.29 Analyse du début du chant kargiraa de la figure 9.31 avec passage de la voix à l'octave inférieure (Son 9.19). Numérotation des harmoniques sur la figure : à gauche : son laryngé normal ; à droite : son en *period-doubling*. Remarquez que le son dédoublé est privé d'énergie sur la fréquence du nouveau fondamental H (1).

Le chanteur commence en mécanisme 1 (voix de poitrine) puis, après un glissando ascendant qui dure à peu près 300 ms, il passe soudainement à l'octave inférieure (flèche rouge sur la figure). Le phénomène est trop rapide pour être décelé à l'oreille. On le lit clairement sur l'analyse spectrographique du transitoire d'attaque (figure 9.29), et très souvent à la fin d'une phrase musicale. En passant ainsi à l'octave inférieure, le chanteur dispose d'une échelle musicale complète (de H8 à H16) dans la zone de F1.

Cette technique permet aux femmes xhosas[15] d'atteindre 110 Hz (la_2) tout en conservant l'intensité d'un son de mécanisme 1 (voir chapitre 6, Son **6.29**). C'est ce que fait aussi le chanteur tibétain présenté au chapitre 6 (§ 2.7.3 figure 6.24 et Son **6.30**) qui, sur un fondamental de 68 Hz ($réb_1$), peut sélectionner l'harmonique 5 à 340 Hz *(fa_3)* par le premier formant d'une voyelle « o » très fermée et très proche du « ou ».

14. L'accolement des bandes ventriculairess ne se produit qu'une période sur deux : la période réelle du son passe au double de celle des cordes vocales et la fréquence fondamentale baisse d'une octave (voire plus quand la période est triple). Voir Bailly, 2008 ; Henrich, 2012.
15. Voir la vidéo réalisée par D. Dargie en 1985, mise en ligne sur « www.youtube.com/watch ?v=MYj-55T6Uzs ». Mme Mbizweni pratique deux techniques différentes. À 4'30, elle chante l'exemple analysé chapitre 6 : *Song iRobhan*e, style umngqokolo ngomqangi.

3.4. Exemples de musique harmonique avec accord d'octave entre F1 et F2

Chant kargiraa. Dans ce chant en technique de *period-doubling* (Son **9.17** et figure 9.30 à gauche), on entend distinctement les voyelles « o » et « a » affectées de différentes nuances. Les deux formants évoluent en parallèle (voir plus haut le schéma de la figure 9.26), ce qui enrichit considérablement la mélodie et lui donne plus d'intensité.

Son 9.17 (19″)

Chant harmonique kargiraa (period-doubling, voir page 472)

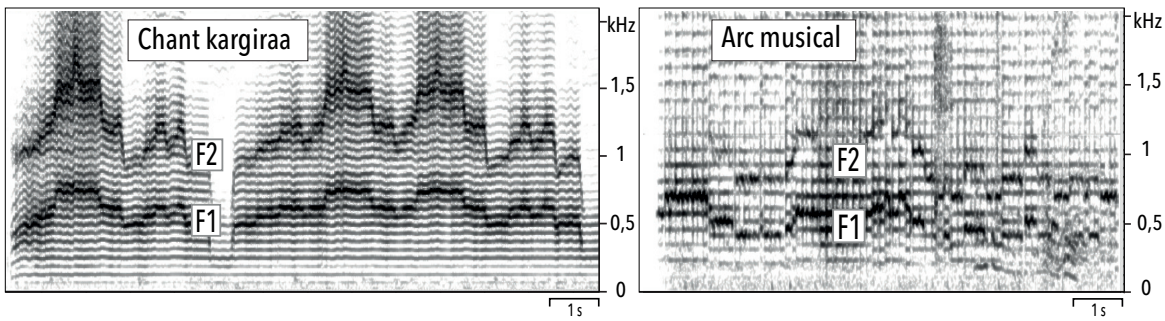


Figure 9.30 Deux exemples de musique harmonique dont la mélodie est réalisée en octave, par les formants 1 et 2 évoluant parallèlement. À gauche : chant harmonique de style kargiraa ; fondamental si_0 = 61,7 Hz en technique vocale de dédoublement. Le chanteur explore la totalité des voyelles sur l'axe du graphique de la figure 9.31 (chant tuva d'Asie centrale). À droite : jeu d'arc musical avec deux fondamentaux : 102 Hz (lab_1) et 114 Hz (sib_1) et de nombreux passages en octave.

Arc musical. L'accordage en octave de F1 et F2 se rencontre aussi à l'arc musical (Son **9.18** et figure 9.30 à droite)[16] mais, dans cet exemple instrumental, la mélodie prend le pas sur l'audition de voyelles, plutôt difficiles à entendre.

Son 9.18 (15″)

Arc musical, Gabon (voir page 472)

Chant tuva octaviant avec ornementation vocalique. Le chant harmonique avec formants en octave, qui n'est possible qu'avec les voyelles « ou », « o », « a », peut paraître limité. Or le Son **9.19**, déjà présenté au chapitre 6 (§ 2.7.3, figure 6.21), offre un bel exemple d'ornementation par emprunt aux deux voyelles « é », « è ».

La figure 9.31 présente l'analyse sonagraphique de ce chant ainsi que les notations musicale et phonétique. Comme précédemment, la ligne mélodique est due aux voyelles comprises entre « o » (ouvert) et « a » (ouvert) dont le premier formant passe du do_4 au sol_4. Nous avons placé sur le graphique de la figure 9.32 les coordonnées formantiques des degrés de la mélodie harmonique sur do_1. Les ornements qui rythment la ligne mélodique supérieure sont réalisés par de brèves incursions empruntées aux diphtongues « ié » et « iè ». Celles-ci provoquent une montée rapide de F2 alors que F1 reste accordé sur l'harmonique de la mélodie principale, de H8 à H12. Pour le la_3, le chanteur baisse passagèrement la fréquence fondamentale en conservant l'accord sur H8 et utilise la voyelle « é ».

Son 9.19 (25″)

Chant harmonique kargiraa (period-doubling) ; F1 et F2 en octaves avec ornementations

16. L'absence de formants dans l'aigu du spectre et surtout les percussions de la corde informent perceptivement qu'il ne peut s'agir d'un son vocal.

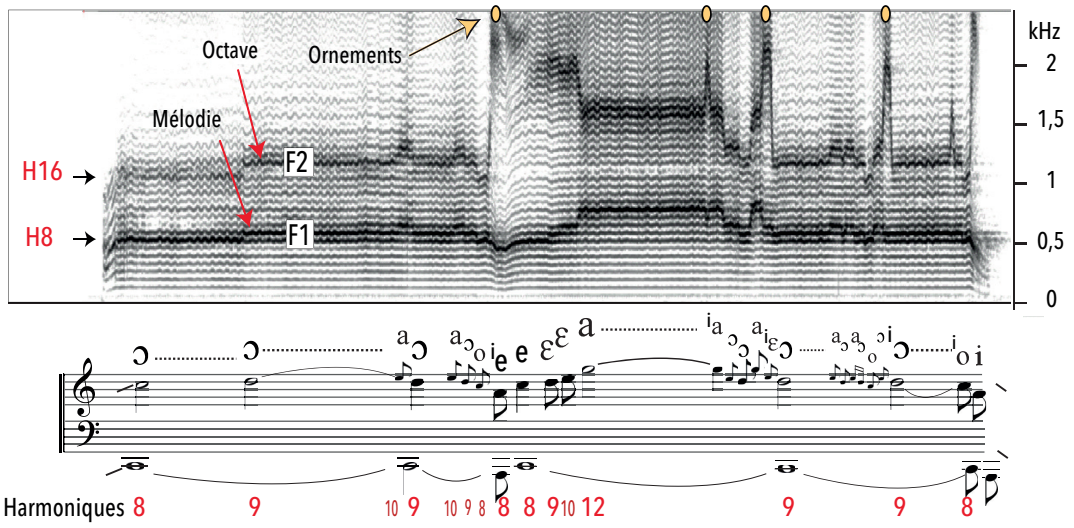


Figure 9.31 Analyse d'un chant tuva kargiraa en technique octaviante (voir la figure 9.29). L'artiste oppose deux groupes de voyelles : les voyelles comprises entre « o » et « a » pour une réalisation extrêmement précise de la mélodie en octaves, et les voyelles antérieures « é » et « è », dont le deuxième formant (F2) est beaucoup plus aigu. Les glissandos spectraux de F2 rythment le chant.

Son 9.19 (25")

*Chant harmonique kargiraa (*period-doubling*) ; F1 et F2 en octaves avec ornementations*

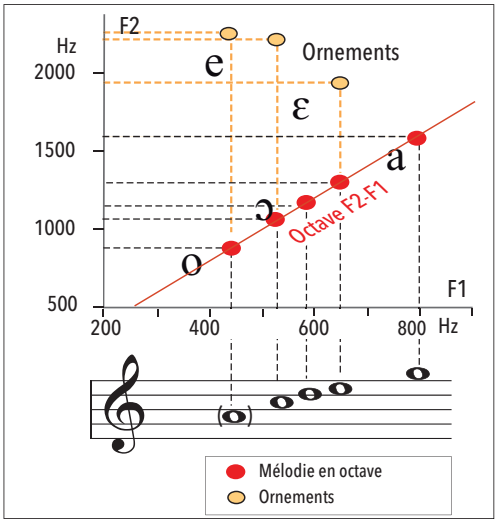


Figure 9.32 Correspondance entre les notes et les voyelles pour le chant analysé figure 9.31.

Les exemples que nous avons présentés sont d'une virtuosité vocale étonnante, tant par le degré de sélectivité harmonique que par la technique d'accordage des formants à l'octave. La transcription phonétique des voyelles proposée est discutable, car l'interprétation perceptive change selon les langues. Les signes phonétiques placés sur les graphiques ont seulement pour fonction d'indiquer des zones laissant une assez grande liberté d'ajustement des résonances.

3.5. La quintina des chanteurs sardes : une expérience perceptive étonnante

Pour clore ce chapitre voix et perception, nous présentons un exemple complexe d'ajustement vocal qui conduit à l'émergence d'une voix virtuelle. Ce phénomène, particulièrement réputé dans le village de Castelsardo, a fait l'objet d'une étude ethnographique et musicologique par B. Lortat-Jacob (1998). L'étude acoustique que nous avons développée au LAM (Castellengo & coll., 2001) nous a permis de mettre en évidence les données cognitives sous-jacentes qui expliquent cette illusion, et d'en apporter la preuve par l'analyse synthèse. En voici la présentation.

Il s'agit de chants polyphoniques pour quatre voix d'hommes nommées, du grave à l'aigu, *bassu*, *contra*, *bogi*, *falzittu*. L'harmonie en est le plus souvent un accord parfait disposé selon la figure 9.33. Dans notre exemple, le bassu entonne seul sur la syllabe « iè » de *Jesus*, puis les trois autres chanteurs entrent en s'ajustant réciproquement au cours d'un portando ascendant. La fréquence fondamentale de l'entrée du bassu est proche du lab_1 (102 Hz)[17].

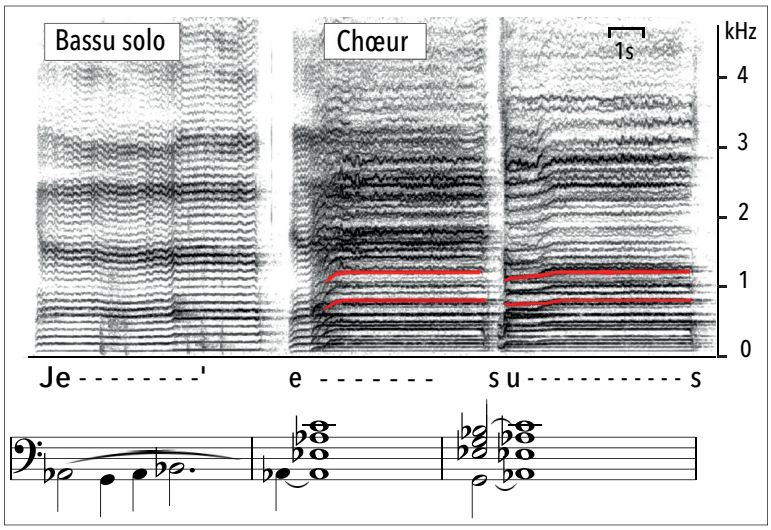


Figure 9.33 Analyse sonagraphique du début de l'exemple musical 9.20. Remarquez la richesse harmonique des voix (on compte plus de 40 harmoniques sur la voix de bassu) et la netteté des formants de la voyelle "è" du bassu. Les tracés en rouge correspondent aux deux composantes filtrées dans le Son 9.21.

À la première écoute (Son **9.20**), l'auditeur est saisi par la plénitude de l'accord et par la richesse harmonique des voix. À la deuxième écoute, il peut explorer une à une les différentes parties chantées et s'étonner de la présence d'une cinquième voix planant au-dessus des quatre chanteurs. Pour certains auditeurs, il faut s'y reprendre à plusieurs fois, mais, lorsqu'elle est repérée, cette voix qui n'est pas comme les autres s'impose avec évidence.

Son 9.20 (42")

Début du chant Jesu *de Castelsardo (Sardaigne)*

17. Il est tentant de simplifier la notation musicale en assimilant la note de départ à un sol_1. Mais, d'une part, la suite du chant se développe principalement (et se termine) sur le *sol*. D'autre part, la hauteur exacte prend sens par le fait que les phénomènes pertinents du point de vue perceptif concernent la zone des harmoniques 8 à 16 de la fréquence fondamentale. Le moindre décalage de la note basse est amplifié sur les harmoniques et l'effet recherché peut s'évanouir. On comprend que les chanteurs soient très attentifs à la hauteur du son de départ.

Pour rendre sensible le phénomène, nous avons utilisé les ressources de l'analyse synthèse[18] qui permet d'extraire ou de supprimer sélectivement des composantes sans altérer la qualité du son. Voici un exemple.

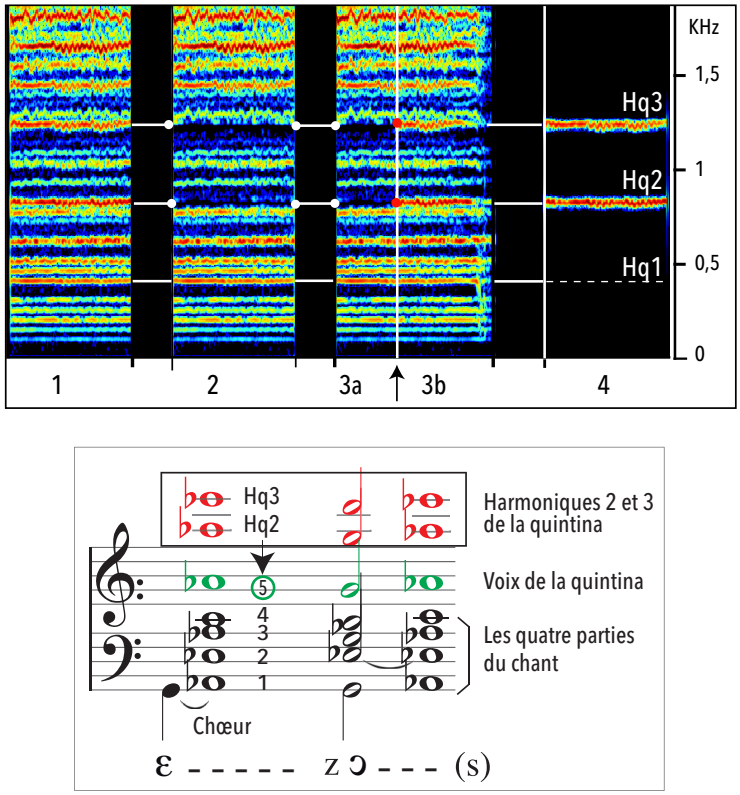


Figure 9.34 Mise en évidence des deux harmoniques générant la voix de quintina. En haut : analyse des différentes présentations du son 9.21. 1 : l'accord extrait du chant ; 2 : l'accord privé de deux harmoniques (Hq2 et Hq3) ; 3 : l'accord filtré (3a) dans lequel les harmoniques sont réintroduits (3b) ; 4 : les deux harmoniques Hq2 et Hq3 entendus isolément. En bas : notation musicale des accords, de la note de la quintina en vert et des deux harmoniques qui la génèrent.

Son 9.21 (18")

Expérience du filtrage sélectif des deux composantes qui créent la quintina (voir détail page 473)

Isolons l'accord qui termine la première phrase du chœur. Le Son **9.21** débute par cet accord (figure 9.34, 1). En (2) l'accord est répété, privé des deux composantes (en rouge sur la figure 9.33) qui génèrent la quintina entendue. Dans la troisième partie, on entend à nouveau l'accord filtré (3a) dans lequel sont réintroduites les deux composantes (3b). En dernier lieu (4), on entend les deux composantes isolées (notées Hq2 et Hq3) qui ont été extraites du son global : la quintina réapparaît, et même une voyelle : « a ». L'effet est saisissant. On remarquera par ailleurs que l'harmonique dont la fréquence correspond à la note fondamentale de la quintina (Hq1) ne joue aucun rôle dans la perception de la cinquième voix : ni dans le son filtré ni dans l'écoute de (4) qui ne comprend que les harmoniques Hq2 et Hq3.

18. Logiciel Audiosculpt développé à l'Ircam.

Cet exemple de perception d'une « voix qui n'existe pas » étonne d'autant plus que le spectre harmonique est extrêmement dense et que deux harmoniques suffisent pour faire émerger une voix supplémentaire. Plusieurs explications ont été proposées, mais seule la synthèse sonore peut confirmer la validité perceptive de l'interprétation que nous proposons.

Pour expliquer l'émergence perceptive de la quintina, il faut rassembler plusieurs des données vues dans les chapitres précédents.

1/ Rappelons tout d'abord que la sensation de hauteur d'une voix est double : elle implique la fréquence fondamentale et la hauteur spectrale formantique (voir chapitres 4 et 6).

2/ Nous avons vu aussi que l'attribution d'une hauteur dite virtuelle se produit le plus souvent lorsque des composantes en rapport de quinte ou de quarte interagissent (voir chapitre 6, figure 6.30).

3/ Nous avons remarqué à plusieurs reprises que les composantes concernées doivent être situées dans la « zone d'écoute dominante », entre 800 et 1600 Hz (voir chapitre 4). Ce point est capital. Les deux composantes filtrées dans le Son **9.21**, lab_4 et mib_5, qui sont interprétées comme les harmoniques 2 et 3 (Hq2 et Hq3) de la voix de quintina, lab_3 (Hq1), sont pleinement dans cette zone dominante pour la perception d'une hauteur virtuelle (voir chapitre 3, encadré page 116, et le chapitre 6, § 3.5.1).

4/ Lorsqu'il s'agit d'une voix, les composantes qui génèrent la hauteur dite virtuelle sont les harmoniques de la voix, renforcés par les formants vocaliques, comme nous l'avons exposé tout au long de ce chapitre.

5. Enfin nous avons vu dans ce chapitre que la réalisation d'un intervalle musical entre les formants – dans cet exemple, il s'agit de la quinte et non plus de l'octave – impose la sélection de certaines voyelles. Ici, le texte du *Jesu* n'est pas véritablement un obstacle puisqu'il est connu de tous, et peut donc subir des transformations vocaliques.

B. Lortat-Jacob (*op. cit.* p. 129) écrit que certaines voyelles, « i » et « ou », sont bannies par les chanteurs et que les autres sont ramenées dans le centre du triangle vocalique (voir figure 9.35). Mais, étant donné que nous avons affaire à un chant polyphonique dont les voix – intimement mêlées – possèdent plusieurs harmoniques communs, et que les voyelles sont différemment distribuées entre les voix, il est difficile de s'avancer plus loin dans l'interprétation du lien entre formants et harmoniques dominants et, à plus forte raison, de tenter d'attribuer l'effet « quintina » à une voix particulière. Seul un enregistrement captant séparément chacune des parties sur des pistes indépendantes permettrait de démêler l'écheveau. Remarquons toutefois que deux groupes de voyelles dominent l'ensemble du chant : le groupe « é-è » et le groupe « a-o ». Avec le groupe « é-è », la zone spectrale 800-1 600 est dégagée, ce qui favorise l'audition des harmoniques de la quintina. Avec le groupe « a-o », les harmoniques de la quintina sont directement renforcés par les formants vocaliques présents dans cette zone, ce qui est le cas de l'exemple sonore que nous avons analysé (voyelle « o » ouvert).

pour le *falzittu*, *i* devient *é*, et *u* devient ɔ ;
pour la *bogi*, *i* devient *è*, et *u* devient ɔ ;
pour le *contra*, *i* devient *è* et *u* devient ɔ ;
pour le *bassu*, *i* devient *ö*, et *u* devient *o*.

Figure 9.35 Ajustement des voyelles par les chanteurs sardes de la quintina.
Lortat-Jacob, B., 1998, page 129.

Son 9.22 (6")

Un accord isolé, répété deux fois : repérer la voix supérieure

Son 9.23 (15")

Exemple de suppléance cognitive dans la phrase complète

6/ Il faut enfin mentionner la suppléance cognitive qui permet d'assurer la continuité mélodique même lorsque la note entendue n'existe pas ou est faible (voir chapitre 6, § 3.5.2 et 3.5.3). Le Son **9.22** en donne un bon exemple. On entend un accord isolé dont la note supérieure est un *ré* (voix du falzittu) pour la plupart des auditeurs. Inséré dans la mélodie (troisième accord du Son **9.23**), le même accord fait entendre le son aigu de la voix de quintina qui est *fa*.

L'émergence d'une cinquième voix, la quintina, résulte donc d'un patient travail d'ajustement vocalique et d'écoute réciproque de la part des chanteurs, ainsi que d'une subtile adaptation aux différentes configurations de l'harmonie (accords parfaits majeur, mineur, accords de quarte et sixte). Un tel art ne prend sens que dans un contexte culturel donné et à un moment précis de l'année : la Semaine sainte. Toutes ces données, essentielles pour comprendre le chant et sa réalisation, sont amplement détaillées dans l'ouvrage que Bernard Lortat-Jacob a consacré aux *Chants de Passion*. Voir également le site : http://crem-cnrs.fr/clefs-ecoute/animations/quintina/seq1.html.

Telle que nous l'avons développée au LAM, cette recherche représente un magnifique exemple faisant appel à plusieurs des particularités de la perception humaine présentées dans cet ouvrage.

3.6. Conclusion

Parmi toutes les sources sonores, la voix humaine est la plus importante perceptivement et la plus riche en possibilités. Adaptable, configurable et se transformant sans cesse, la voix donne lieu à une multiplicité d'usages et donc d'écoutes. L'unique point que nous avons abordé ici, celui des relations entre la fréquence fondamentale du son laryngé et celles des résonateurs pharyngo-buccaux, témoigne de la complexité et de la richesse de ce domaine d'étude.

Conditionnés par l'anatomie et la physiologie de la tête, les mouvements des organes phonatoires agissent sur les modifications couplées des divers résonateurs, de sorte que les premiers formants restent dans des rapports similaires, malgré les différences morphologiques dues à l'âge ou au sexe. Face à la variabilité des réalisations acoustiques, les auditeurs développent deux types de stratégies perceptives évoquées à plusieurs reprises dans les chapitres précédents. D'une part, la recherche de relations stables, celles que l'on peut catégoriser de façon significative en ignorant les différences de présentation, d'autre part, l'attention extrême portée aux variations les plus minimes qui sont alors interprétées en termes de différences qualitatives. Ce que nous appelons « voyelles » se prête admirablement bien à des écoutes multiples. Les voyelles sont organisées selon une constellation sonore de dimensions globales variables, mais dont les éléments sont dans des rapports internes très précis. Selon l'auditeur, la langue, le style musical, les variations de ces rapports font sens ou sont interprétées comme changements de qualité vocale.

La voix chantée est diversement travaillée selon les cultures. Du seul point de vue de la perception de la hauteur, un chanteur négocie en permanence les rapports entre la fréquence fondamentale de la source laryngée et celles des résonances des cavités, pour répondre aux contraintes parfois contradictoires de l'intelligibilité du texte et de la qualité vocale. Nous n'avons présenté que deux exemples de chant lyrique, le sujet nécessiterait une étude systématique prenant en compte les différentes catégories vocales et les styles de chant : chanson, mélodie, opéra.

L'accord spécifique entre un harmonique et une fréquence de résonance (*formant-tuning*) est un effet recherché dans certains styles musicaux et en particulier sur de longues tenues, comme dans certains chants rituels tibétains, toujours – semble-t-il – sur la voyelle « o » très fermée. La même voyelle sert un tout autre effet dans l'exemple du *Pasi but but* des Bunun, qui par ailleurs témoigne d'une véritable *registration* spectrale par la distribution des voyelles aux quatre voix de la polyphonie.

Dans d'autres cultures, particulièrement en Asie centrale, les fréquences de résonance des cavités sont travaillées de façon sélective pour donner à entendre une mélodie fournie par les harmoniques du son glottique. Les deux résonances F1 et F2 sont présentes mais la focalisation se fait tantôt sur l'une, tantôt sur l'autre. Les chants de F2 ont été analysés au chapitre 6, § 2.7.2 et au chapitre 8, § 1.3.1. Nous avons plus longuement développé dans ce chapitre les chants de F1 qui sont associés à une technique vocale particulière de dédoublement de la fréquence fondamentale de la voix (*period-doubling*), de sorte que les harmoniques au voisinage de la zone de résonance utilisée soient en nombre suffisant. Les exemples analysés sont majoritairement masculins (kargiraa). Cependant, la technique est aussi utilisée par les femmes xhosas d'Afrique du Sud, présentées au chapitre 6, qui pratiquent également l'accord des deux formants à l'octave sur deux notes fondamentales alternées, à l'imitation du jeu de l'arc musical.

La richesse et la diversité des techniques vocales développées dans les cultures traditionnelles ne cesse d'étonner et tout près de nous celle des Sardes qui, si elle fascine les auditeurs à la première écoute, fascine tout autant le chercheur qui s'attache à en démêler les ressorts perceptifs.

Plusieurs des exemples sonores présentés dans ce chapitre sont inclus dans une publication aujourd'hui épuisée[19]. On peut de nouveau y accéder en ligne à l'adresse : archives.crem-cnrs.fr.

4. Les sons du chapitre 9

Son 9.1 – Typologie sonore. Cinq sons de l'environnement, chacun étant représentatif d'une catégorie acoustique. 1/ débouchage (bouteille de vin) ; 2/ grincement (chaise) ; 3/ souffle sur fleurs séchées ; 4/ klaxon : 5/ ondes Martenot[20] [M. C.]

Son 9.2 – Typologie phonatoire. Cinq sons vocaux relevant des catégories acoustiques de l'exemple précédent et un sixième, spécifique de l'appareil phonatoire. [M. C.]

Son 9.3 – Émission alternée sur la voyelle « e » : voix chuchotée, voisée, chuchotée. Voix féminine. [M. C.]

Son 9.4 – Signal électroglottographique (EGG) capté au niveau du larynx pendant le chant. Chanteur baryton, B. Chuberre ; thèse N. Henrich, 2001. [Archives LAM]

Son 9.5 – Signal sonore capté par un microphone placé face à la bouche du chanteur, simultanément avec le signal EGG du Son 9.4 ; début de l'*Ave Maria* de Gounod. [*Ibid.*]

Son 9.6 – Glissando ascendant-descendant rapide chanté avec trois voyelles, « i », « a », « ou ». Voix masculine, S. Lamesch. [M. C.]

19. *Les Voix du monde*, coffret de 3 CD édité par le CNRS et le laboratoire d'ethnomusicologie du musée de l'Homme.
20. Sons 1 et 4 extraits de la collection A (Son 4.4 ; n°s 9 et 17) ; sons 2 et 3 extraits de la collection B (Son 4.5 ; n°s 28 et 13).

Son 9.7 – Voyelles « i, a, ou » articulées avec une voix très grave en mécanisme 0 (fry). Voix masculine, S. Lamesch. [M. C.]

Son 9.8 – Voyelles « i, a, ou » articulées avec une voix très grave en mécanisme 0 (fry). Voix féminine, I. Wollman. [M. C.]

Son 9.9 – Séquence de 11 voyelles enchaînées en mécanisme 0 (fry). Voix masculine, S. Lamesch. [M. C.]

Son 9.10 – Une soprano chante une courte phrase à différentes hauteurs de sa voix, en commençant sur un son très aigu (1) : *ré*$_5$ (1176 Hz). La phrase qu'elle articule est incompréhensible. Elle poursuit en descendant (2) : *la*$_4$ (880 Hz) ; pouvez-vous noter ce qu'elle dit ? Sur le son suivant (3) : *ré*$_4$ (588 Hz), la phrase est parfaitement intelligible. La chanteuse continue à descendre (4) : *la*$_3$ (440 Hz), et (5) : *ré*$_3$ (294 Hz). Lors de la remontée inverse, vous noterez que maintenant vous comprenez aisément la phrase chantée sur le son (2), *la*$_4$, parce que vous la connaissez. Chanteuse Mme Ponthié, 1969. [Archives LAM]

Son 9.11 – Gamme diatonique chantée sur la voyelle « é » de *do*$_2$ à *do*$_3$ (diapason haut). Baryton, B. Chuberre ; thèse N. Henrich, 2001. [Archives LAM]

Son 9.12 – Trois changements de qualité vocale sur la voyelle « a » : note *do*$_4$. Annonces faites par la chanteuse avant chaque son : « émission normale ; son couvert ; et son clair ». Soprano C. Herzog, 1979. [Archives LAM]

Son 9.13 – Trois changements de qualité vocale sur la voyelle « a » : note *ré*$_4$. Annonces faites par la chanteuse avant chaque émission : « son normal ; son couvert ; son clair ». [*Ibid.*]

Son 9.14 – Chant *Pasi but but* des Bunun de Naifubo (Taïwan, 1987). La voix supérieure (voix 1) commence faiblement sur un *sol*$_2$ (193Hz) puis monte peu à peu jusqu'au do$_3$ tout en conservant la voyelle « o ». Les autres voix entrent tour à tour en s'ajustant sur la voix supérieure qui s'élève : voix 2, à la tierce mineure inférieure de 1 (voyelle « é ») ; voix 3, à la quarte inférieure de 1 (voyelle « é ») ; voix 4, à la quinte inférieure de 1 (voyelle « i »). Au bout du 6e cycle, le chant se termine sur la quinte finale. Remarquez l'emploi de voyelles distinctes selon les voix. Document Wu Rung Shun, thèse, 1996. [Archives CREM et archives LAM]

Son 9.15 – Chant de style gidayu (Japon) illustrant le croisement complexe des hauteurs tonale (fréquence laryngée) et spectrale (variations formantiques dues aux voyelles). Très court extrait (7 '') suivi du même ralenti deux fois. [Référence non identifiée]

Son 9.16 – Deux exemples de chant harmonique sur la même note fondamentale, *do*$_2$. Au début, le chanteur utilise la résonance vocalique de F2 (deuxième formant) puis, toujours sur la même fondamentale, celle de F1 (premier formant). Chanteur Trân Quang Hai, 1985. [Archives LAM]

Son 9.17 – Chant harmonique tuva de style kargiraa *Artii -Sayir* utilisant F1 et F2 en octaves. Noter l'ajustement des voyelles : de « o » (ouvert) vers « a » (ouvert) dans le sens mélodique ascendant. Emission vocale en *period-doubling*. CD *Tuva, Voices from the center of Asia,* plage 18, Smithsonian/Folkways, 1990. [SF 40017]

Son 9.18 – Court extrait de jeu à l'arc musical (Gabon). La qualité sonore particulière de cet extrait tient au fait que la mélodie de l'arc est ici contrôlée par les formants F1 et F2 en octave. Il est toutefois difficile d'y percevoir des voyelles en raison de la dissociation cognitive entre la source (un arc) et des voyelles habituellement portées par une voix humaine. Pierre Sallée, 1965, Missoko. [Archives CREM, CD 23, collection CNRSMH_I_2007_001]

Son 9.19 – Chant harmonique tuva de style kargiraa dans lequel le chanteur ajuste F1 et F2 en octaves (voyelles « a » et « o »), et produit de subtils ornements spectraux et rythmiques en utilisant les voyelles « i » et « é ». Émission vocale en *period-doubling*. CD *Tuva, Voices from the center of Asia,* plage 1, Smithsonian/Folkways,1990. [SF 40017]

Son 9.20 – Chant de Sardaigne (début du *Jesu)* dans lequel on entend la « quintina ». Chanteurs de Castelsardo. *Les Voix du monde*, CD 3-18, Harmonia Mundi. Document B. Lortat-Jacob, 1995. Chant complet accessible sur le site « crem.telemeta.org/archives/collections/CNRSMH_E_1996_013_001 ». [Archives CREM]

Son 9.21 – Expériences de filtrage pour faire entendre la quintina sur le 2e accord du chant précédent. On entend successivement quatre courtes séquences : 1) un accord ; 2) l'accord filtré (sans les deux composantes qui produisent la quintina) ; 3a) l'accord filtré et (3b) la ré introduction – à la 12e seconde – des deux composantes qui génèrent la voix de quintina ; 4) les deux composantes isolées qui créent la voix de quintina dont elles sont les harmoniques 2 et 3 (Hq2 et Hq3 de la figure 9.34). [M. C.]

Son 9.22 – Écoute d'un accord isolé (court), répété deux fois. Repérer et mémoriser la « note » supérieure de l'accord (un *ré* pour beaucoup d'auditeurs). [M. C.]

Son 9.23 – Phrase extraite du chant du *Jesu*. La quintina évolue parallèlement à la basse, en tierce, et l'on entend à la partie supérieure: *ré, mi, fa, mi, ré, do*#. Or le son le plus aigu, *fa*, est produit par l'accord de l'exemple 9.22. En contexte, le *fa* aigu s'impose pour assurer la continuité mélodique du chant. Ce phénomène de suppléance cognitive a été décrit par les Gestaltistes. [M. C.]

ANNEXES

ANNEXE A
Conventions de notation musicale 477

ANNEXE B
Intervalles musicaux 479

ANNEXE C
Pratique du lecteur musicien 485

ANNEXE D
Textes 487

ANNEXE E
Visualisation des phénomènes vibratoires 489

ANNEXE F
Bibliographie 493

ANNEXE G
Glossaire 515

ANNEXE H
Contenu du DVD-Rom d'accompagnement 531

ANNEXE A

Conventions de notation musicale

La désignation des degrés de l'échelle chromatique et celle des octaves dans lesquelles se situent les sons diffèrent selon les traditions musicales. En Europe coexistent trois systèmes qui relèvent des grandes traditions culturelles : *latine* (Italie, France, Espagne), *anglo-saxonne* (Angleterre, États-Unis), et *germanique* (Allemagne, Europe centrale et Europe du Nord). La diffusion d'instruments électroniques et de logiciels ainsi que la norme MIDI tendent aujourd'hui à généraliser le système anglo-saxon. Cependant les instrumentistes, les facteurs d'instruments, de même que les chercheurs musicologues, adoptent toujours les dénominations en usage dans chacune des cultures. Il importe donc de les connaître toutes pour lire les textes anciens et modernes. Enfin, pour éviter toute ambiguïté, il est recommandé de toujours associer la fréquence à la dénomination des sons.

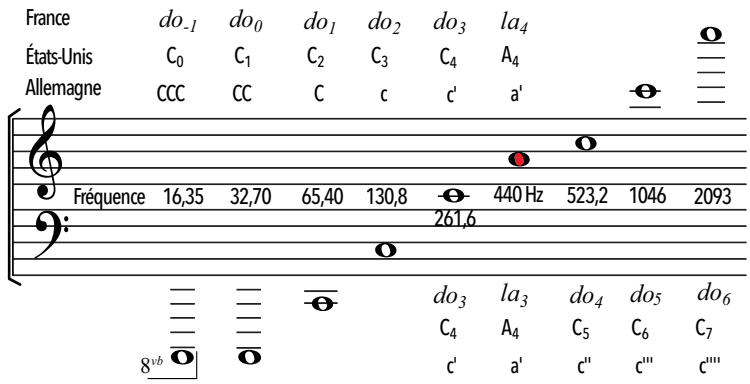


Figure A.1 Numérotation des octaves en usage dans différentes cultures. Fréquences des « *do* » dans le système du tempérament égal (référence la_3 = 440 Hz)

La figure A.1 indique la numérotation des *do.* Originellement fondée sur la tessiture de la voix humaine (do_1-do_5), la numérotation des octaves s'est progressivement étendue vers les graves – octaves 0 et -1 en notation latine – ce qui a entraîné le décalage d'une unité dans la notation anglo-saxonne pour éliminer le signe négatif. Tous les sons inclus dans l'intervalle ascendant *do-si* portent le même numéro. Il faut noter aussi que plusieurs auteurs allemands écrivent « a_1 » pour « *a'* ». Le son du diapason (440 Hz) s'écrit donc, selon les conventions : la_3, A4 ou a_1. Le respect de la casse (majuscule ou minuscule) est impératif dans la notation germanique.

La dénomination des sons de la gamme et celle des sons altérés diffèrent également selon les trois traditions. La figure A.2 montre la correspondance entre la notation musicale sur portée et les noms des degrés d'une échelle chromatique en usage dans les trois types de notation mentionnées. On remarquera certaines particularités dans la désignation des degrés diésés et bémolisés, en particulier l'écriture du *si* bémol et du *si* bécarre en anglais et en allemand.

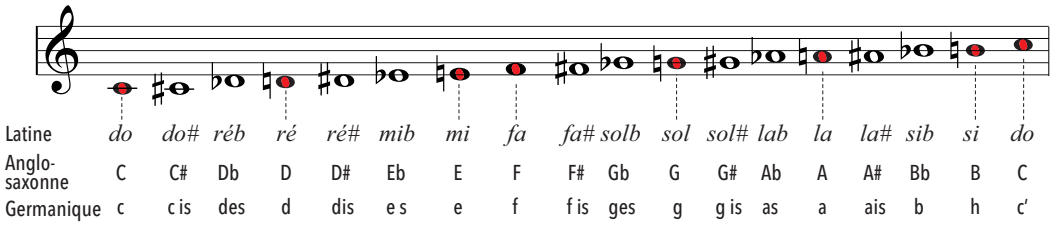


Figure A.2 Conventions de désignation des sons de l'échelle chromatique dans les trois traditions européennes.

ANNEXE B

Intervalles musicaux

L'estimation et la mesure des intervalles musicaux est présentée et développée au chapitre 8, encadrés pages 405 et 406. Nous donnons ci-après une documentation complémentaire.

1. Tableau des intervalles et de leurs mesures

Mesure d'un intervalle en unités logarithmiques : savarts ou cents (remarque : le résultat est toujours une approximation.)

Il faut partir du rapport des fréquences ou du rapport entre les numéros d'harmoniques (colonne : Rapport harmonique). Pour les cents, on convertit le résultat en $\log_2$ et on le multiplie par 1200. Pour les savarts, on convertit le résultat en $\log_{10}$ et on le multiplie par 1000.

Exemple pour la tierce majeure pure, entre les harmoniques 4 et 5. On a 5/4 = 1,25

- $\log_2 (1{,}25) \times 1200 = 386{,}31$ cents
- $\log_{10} (1{,}25) \times 1000 = 96$ savarts

Quelques valeurs d'intervalles à retenir :

Un ton tempéré = 200 cents ou 50 savarts.

Un quart de ton tempéré = 50 cents ou 12,5 savarts.

Lorsqu'on connaît les fréquences de deux sons proches, on peut estimer rapidement l'intervalle qui les sépare sur la base du rapport numérique des fréquences. Exemple : la_3 et $sol\#_3$.

Rapport des fréquences = 440/415 = 1,06 ou 6 %

Un rapport de 6 % caractérise le demi-ton du tempérament égal (en toute rigueur, c'est 1,05946, la valeur de la racine douzième de 2). De la même façon, on peut estimer le ton tempéré (1,122 soit 12 %) et le quart de ton tempéré (3 %).

Les quotients des fractions simples comme celles de la quinte (3/2 = 1,5), de la quarte (4/3 = 1,33) et de la tierce majeure (5/4 = 1,25) sont aussi à mémoriser pour s'y retrouver rapidement sans table numérique ni ordinateur.

Intervalle	Rapport harmonique	Valeur numérique	Mesure en cents	Mesure en savarts
Octave	2/1	2	1200	301
Septième Maj t		1,887	1100	275
Septième Maj	15/8	1,875	1088,2	273
Septième min 2	9/5	1,8	1017,6	255,27
Septième min t		1,78	1000	250
Septième min 1	7/4	1,75	968,82	243,03
Sixte Maj 2 (ou 7e)	12/7	1,71	933,12	234
Sixte min t		1,68	900	225
Sixte Maj 1	5/3	1,66	884,35	221,84
Sixte min	8/5	1,6	813,68	204,11
Sixte min t		1,587	800	200
Quinte	3/2	1,5	701,95	176
Quinte t		1,498	700	175
Triton Gd	36/25	1,44	630	158
Triton Maj	10/7	1,428	617,48	154,9
Triton t		1,414	600	150
Triton min	7/5	1,4	582,51	146,12
Quarte t		1,334	500	125
Quarte	4/3	1,33	498	124,9
Tierce Gde	9/7	1,28	435,08	109,14
Tierce Maj t		1,259	400	100
Tierce Maj	5/4	1,25	386,31	96,9
Tierce min 2	6/5	1,20	315,64	79,1
Tierce min t		1,189	300	75
Tierce min 1	7/6	1,166	266,87	66,9
Ton Gd	8/7	1,14	231,17	57,99
Ton Maj	9/8	1,125	203,91	51,15
Ton t		1,122	200	50
Ton min	10/9	1,111	182,4	45,75
Demi-ton Maj	16/15	1,066	111,7	28
Demi-ton t		1,059	100	25
Demi-ton ± égal	18/17	1,0588	98,95	24,82
Demi-ton min	25/24	1,041	70,67	17,72

Figure B.1 Tableau comparatif des valeurs numériques des principaux intervalles musicaux. La colonne 1 (texte en rouge) concerne exclusivement la série harmonique dont les quatre premiers intervalles (octave, quarte, quinte, tierce majeure) sont sur fond beige. Les intervalles de la gamme chromatique en tempérament égal sont sur fond bleu. Les colonnes 4 et 5 donnent la valeur de ces intervalles en unités logarithmiques cents et savarts.

2. Calcul des commas

La fonction musicale des commas a été présentée au chapitre 8, § 3.1.3. Nous développons ci-dessous le détail du calcul qui permet d'obtenir leurs valeurs en cents. Pour effectuer aisément ces opérations, il faut savoir manipuler la multiplication et la divisions des rapports et ne pas perdre de vue l'octave dans laquelle se situent les sons.

2.1. Calcul du comma syntonique

Méthode par quintes ascendantes

Partant du do_3 nous aboutissons à un mi_5 qu'il faut comparer avec l'harmonique 5 du do_3 Nous utiliserons l'indice « q » pour le *mi* issu des 4 quintes, et l'indice « h » pour le *mi* harmonique en rapport de tierce pure avec le *do*.

$mi_{5q}/do_3 = (mi_{5q}/la_4) \times (la_4/ré_4) \times (ré_4/sol_3) \times (sol_3/do_3)$; ce produit de 4 quintes s'écrit $(3/2)^4 = 81/16 = 5{,}0625$.

L'intervalle mi_{5q}/do_3 vaut 5,0625 ; l'intervalle mi_{5h}/do_3 vaut 5.

Le mi_{5q} généré par 4 quintes est plus haut que le mi_{5h} harmonique 5 du do_3.

L'intervalle entre les deux *mi* est donné par le rapport 5,0625/5. Pour convertir le résultat en cents, on prend le $\log_2$ du rapport que l'on multiplie par 1200.

Méthode par quintes ascendantes et quartes descendantes

On pose : $mi_{3q}/do_3 = (mi_{3q}/la_3) \times (la_3/ré_3) \times (ré_3/sol_3) \times (sol_3/do_3)$ c'est-à-dire $(3/4) \times (3/2) \times (3/4) \times (3/2) = 81/64 = 1{,}265625$

On a par ailleurs : $mi_{3h}/do_3 = 5/4 = 1{,}25$

Comma syntonique (cents) = $\log_2$ (1,265 625/1,25) × 1 200 = 21,5 cents

Remarque : il est plus précis d'effectuer le calcul du logarithme sur les rapports que sur les résultats arrondis que nous donnons.

2.2. Calcul du comma pythagoricien

Il faut douze quintes et quartes pures alternées pour parvenir à un son voisin de l'octave de départ. Dans notre exemple, on lit 6 quintes ascendantes (3/2) et 6 quartes descendantes (3/4) pour obtenir le *si*#, soit $(3/2)^6 \times (4/3)^6$, ce qui donne un nombre supérieur à 2.

On peut aussi procéder simplement par tons (3/2) x (3/4) = 9/8 : il y en a 6 soit $(9/8)^6$, ce qui donne (531 441)/(262 144) ≈ 2,027 286. Dans les deux cas on aboutit à un nombre supérieur à 2.

Comma pythagoricien (cents) = $\log_2$ (2,027 286/2) × 1 200 = 32,46 cents

2.3. Calcul du comma enharmonique

Il faut revenir aux tierces pures et comparer une succession de 3 tierces pures à l'octave.

$(5/4)^3 = 1{,}953125$ ce qui est inférieur à 2.

Comma enharmonique (cents) = $\log_2$ (2/1,953 125) × 1 200 = 41,05 cents

soit à peu près le double du comma syntonique.

3. Divers

3.1. Mesure de la fréquence d'un son avec un accordeur

L'accordeur électronique permet de mesurer rapidement l'écart d'un son instrumental donné par rapport à une référence. Les applications disponibles sur smartphones remplacent aisément les appareils dédiés. Ils fournissent le résultat en cents pour un système donné (tempérament égal, tempéraments historiques, modes orientaux), et pour un diapason quelconque. L'accordage est d'autant plus facile que l'instrument possède beaucoup d'énergie dans le fondamental, car c'est sur cette fréquence que s'opère la mesure. Pour certains instruments, il peut subsister un écart notable entre la mesure et la justesse perçue à l'oreille, en particulier dans l'extrême grave (inharmonicité des cordes raides, voir chapitre 6, § 3) et dans l'extrême aigu où des écarts notables existent entre la mesure et la perception de la hauteur musicale (voir chapitre 3, § 2.7.3). Il arrive aussi fréquemment que l'appareil, ou le logiciel, mesure une composante du son autre que le fondamental. L'oreille reste seul juge !

3.2. Trouver la fréquence d'un son dont on connaît l'intervalle par rapport à un autre

Il est commode de se référer à une table numérique dans laquelle on trouve les fréquences de tous les degrés chromatiques dans une octave, par progression de 1 cent (1/100 de demi-ton tempéré). Les tableaux des deux pages suivantes fournissent la fréquence (en Hz) des sons de l'octave 3 pour un diapason de référence la_3 = 440 Hz. Par exemple, la fréquence d'un la_3 situé un quart de ton plus bas est sur la ligne –50 cents, soit 427,47 Hz. Pour les sons des octaves supérieures et inférieures, multiplier (ou diviser) les fréquences par 2 et ses puissances.

4. Fréquences des notes du tempérament égal calculées pour l'octave 3

Fréquence décroissante

Cent	*do*	*do#* *réb*	*ré*	*ré#* *mib*	*mi*	*fa*	*fa#* *solb*	*sol*	*sol#* *lab*	*la*	*la#* *sib*	*si*	Cent
Réf. 0	261,63	277,18	293,66	311,13	329,63	349,23	369,99	392,00	415,30	440,00	466,16	493,88	0
-1	261,47	277,02	293,50	310,95	329,44	349,03	369,78	391,77	415,06	439,75	465,89	493,60	-1
-2	261,32	276,86	293,33	310,77	329,25	348,83	369,57	391,54	414,83	439,49	465,63	493,31	-2
-3	261,17	276,70	293,16	310,59	329,06	348,62	369,35	391,32	414,59	439,24	465,36	493,03	-3
-4	261,02	276,54	292,99	310,41	328,87	348,42	369,14	391,09	414,35	438,98	465,09	492,74	-4
-5	260,87	276,38	292,82	310,23	328,68	348,22	368,93	390,86	414,11	438,73	464,82	492,46	-5
-6	260,72	276,22	292,65	310,05	328,49	348,02	368,71	390,64	413,87	438,48	464,55	492,17	-6
-7	260,57	276,06	292,48	309,87	328,30	347,82	368,50	390,41	413,63	438,22	464,28	491,89	-7
-8	260,42	275,90	292,31	309,69	328,11	347,62	368,29	390,19	413,39	437,97	464,01	491,61	-8
-9	260,27	275,75	292,14	309,51	327,92	347,42	368,08	389,96	413,15	437,72	463,75	491,32	-9
-10	260,12	275,59	291,97	309,34	327,73	347,22	367,86	389,74	412,91	437,47	463,48	491,04	-10
-11	259,97	275,43	291,80	309,16	327,54	347,02	367,65	389,51	412,67	437,21	463,21	490,76	-11
-12	259,82	275,27	291,64	308,98	327,35	346,82	367,44	389,29	412,44	436,96	462,94	490,47	-12
-13	259,67	275,11	291,47	308,80	327,16	346,62	367,23	389,06	412,20	436,71	462,68	490,19	-13
-14	259,52	274,95	291,30	308,62	326,97	346,42	367,01	388,84	411,96	436,46	462,41	489,91	-14
-15	259,37	274,79	291,13	308,44	326,78	346,22	366,80	388,61	411,72	436,20	462,14	489,62	-15
-16	259,22	274,63	290,96	308,26	326,60	346,02	366,59	388,39	411,48	435,95	461,88	489,34	-16
-17	259,07	274,47	290,80	308,09	326,41	345,82	366,38	388,17	411,25	435,70	461,61	489,06	-17
-18	258,92	274,32	290,63	307,91	326,22	345,62	366,17	387,94	411,01	435,45	461,34	488,77	-18
-19	258,77	274,16	290,46	307,73	326,03	345,42	365,96	387,72	410,77	435,20	461,08	488,49	-19
-20	258,62	274,00	290,29	307,55	325,84	345,22	365,74	387,49	410,53	434,95	460,81	488,21	-20
-21	258,47	273,84	290,12	307,38	325,65	345,02	365,53	387,27	410,30	434,70	460,54	487,93	-21
-22	258,32	273,68	289,96	307,20	325,47	344,82	365,32	387,05	410,06	434,44	460,28	487,65	-22
-23	258,17	273,52	289,79	307,02	325,28	344,62	365,11	386,82	409,82	434,19	460,01	487,37	-23
-24	258,02	273,37	289,62	306,84	325,09	344,42	364,90	386,60	409,59	433,94	459,75	487,08	-24
-25	257,87	273,21	289,45	306,67	324,90	344,22	364,69	386,38	409,35	433,69	459,48	486,80	-25
-26	257,73	273,05	289,29	306,49	324,71	344,02	364,48	386,15	409,11	433,44	459,22	486,52	-26
-27	257,58	272,89	289,12	306,31	324,53	343,82	364,27	385,93	408,88	433,19	458,95	486,24	-27
-28	257,43	272,74	288,95	306,14	324,34	343,63	364,06	385,71	408,64	432,94	458,68	485,96	-28
-29	257,28	272,58	288,79	305,96	324,15	343,43	363,85	385,48	408,41	432,69	458,42	485,68	-29
-30	257,13	272,42	288,62	305,78	323,96	343,23	363,64	385,26	408,17	432,44	458,16	485,40	-30
-31	256,98	272,26	288,45	305,61	323,78	343,03	363,43	385,04	407,93	432,19	457,89	485,12	-31
-32	256,83	272,11	288,29	305,43	323,59	342,83	363,22	384,82	407,70	431,94	457,63	484,84	-32
-33	256,69	271,95	288,12	305,25	323,40	342,63	363,01	384,59	407,46	431,69	457,36	484,56	-33
-34	256,54	271,79	287,95	305,08	323,22	342,44	362,80	384,37	407,23	431,44	457,10	484,28	-34
-35	256,39	271,64	287,79	304,90	323,03	342,24	362,59	384,15	406,99	431,19	456,83	484,00	-35
-36	256,24	271,48	287,62	304,72	322,84	342,04	362,38	383,93	406,76	430,94	456,57	483,72	-36
-37	256,09	271,32	287,46	304,55	322,66	341,84	362,17	383,71	406,52	430,70	456,31	483,44	-37
-38	255,95	271,16	287,29	304,37	322,47	341,65	361,96	383,49	406,29	430,45	456,04	483,16	-38
-39	255,80	271,01	287,12	304,20	322,28	341,45	361,75	383,26	406,05	430,20	455,78	482,88	-39
-40	255,65	270,85	286,96	304,02	322,10	341,25	361,54	383,04	405,82	429,95	455,52	482,60	-40
-41	255,50	270,70	286,79	303,85	321,91	341,05	361,33	382,82	405,58	429,70	455,25	482,32	-41
-42	255,35	270,54	286,63	303,67	321,73	340,86	361,13	382,60	405,35	429,45	454,99	482,05	-42
-43	255,21	270,38	286,46	303,49	321,54	340,66	360,92	382,38	405,12	429,21	454,73	481,77	-43
-44	255,06	270,23	286,30	303,32	321,36	340,46	360,71	382,16	404,88	428,96	454,47	481,49	-44
-45	254,91	270,07	286,13	303,14	321,17	340,27	360,50	381,94	404,65	428,71	454,20	481,21	-45
-46	254,77	269,91	285,96	302,97	320,98	340,07	360,29	381,72	404,42	428,46	453,94	480,93	-46
-47	254,62	269,76	285,80	302,79	320,80	339,87	360,08	381,50	404,18	428,22	453,68	480,66	-47
-48	254,47	269,60	285,63	302,62	320,61	339,68	359,88	381,28	403,95	427,97	453,42	480,38	-48
-49	254,32	269,45	285,47	302,44	320,43	339,48	359,67	381,06	403,71	427,72	453,15	480,10	-49
-50	254,18	269,29	285,30	302,27	320,24	339,29	359,46	380,84	403,48	427,47	452,89	479,82	-50
Cent	*do*	*do#* *réb*	*ré*	*ré#* *mib*	*mi*	*fa*	*fa#* *solb*	*sol*	*sol#* *lab*	*la*	*la#* *sib*	*si*	Cent

Figure B.2 Fréquences des sons d'une gamme chromatique de l'octave 3. Tempérament égal (la_3 = 440 Hz). Valeurs décroissantes, par pas de 1 cent.

Fréquence croissante

Cent	*do*	*do# réb*	*ré*	*ré# mib*	*mi*	*fa*	*fa# solb*	*sol*	*sol# lab*	*la*	*la# sib*	*si*	Cent
50	269,29	285,30	302,27	320,24	339,29	359,46	380,84	403,48	427,47	452,89	479,82	508,36	50
49	269,14	285,14	302,10	320,06	339,09	359,25	380,62	403,25	427,23	452,63	479,55	508,06	49
48	268,98	284,98	301,92	319,87	338,89	359,05	380,40	403,02	426,98	452,37	479,27	507,77	48
47	268,83	284,81	301,75	319,69	338,70	358,84	380,18	402,78	426,73	452,11	478,99	507,48	47
46	268,67	284,65	301,57	319,50	338,50	358,63	379,96	402,55	426,49	451,85	478,72	507,18	46
45	268,52	284,48	301,40	319,32	338,31	358,42	379,74	402,32	426,24	451,59	478,44	506,89	45
44	268,36	284,32	301,22	319,14	338,11	358,22	379,52	402,09	426,00	451,33	478,16	506,60	44
43	268,21	284,15	301,05	318,95	337,92	358,01	379,30	401,85	425,75	451,07	477,89	506,30	43
42	268,05	283,99	300,88	318,77	337,72	357,80	379,08	401,62	425,50	450,81	477,61	506,01	42
41	267,90	283,83	300,70	318,58	337,53	357,60	378,86	401,39	425,26	450,54	477,34	505,72	41
40	267,74	283,66	300,53	318,40	337,33	357,39	378,64	401,16	425,01	450,28	477,06	505,43	40
39	267,59	283,50	300,36	318,22	337,14	357,18	378,42	400,93	424,77	450,02	476,78	505,14	39
38	267,43	283,33	300,18	318,03	336,94	356,98	378,21	400,69	424,52	449,76	476,51	504,84	38
37	267,28	283,17	300,01	317,85	336,75	356,77	377,99	400,46	424,28	449,50	476,23	504,55	37
36	267,12	283,01	299,84	317,66	336,55	356,57	377,77	400,23	424,03	449,25	475,96	504,26	36
35	266,97	282,84	299,66	317,48	336,36	356,36	377,55	400,00	423,79	448,99	475,68	503,97	35
34	266,81	282,68	299,49	317,30	336,17	356,15	377,33	399,77	423,54	448,73	475,41	503,68	34
33	266,66	282,52	299,32	317,11	335,97	355,95	377,11	399,54	423,30	448,47	475,13	503,39	33
32	266,51	282,35	299,14	316,93	335,78	355,74	376,90	399,31	423,05	448,21	474,86	503,10	32
31	266,35	282,19	298,97	316,75	335,58	355,54	376,68	399,08	422,81	447,95	474,59	502,81	31
30	266,20	282,03	298,80	316,57	335,39	355,33	376,46	398,85	422,56	447,69	474,31	502,52	30
29	266,04	281,86	298,63	316,38	335,20	355,13	376,24	398,62	422,32	447,43	474,04	502,23	29
28	265,89	281,70	298,45	316,20	335,00	354,92	376,03	398,39	422,08	447,17	473,76	501,94	28
27	265,74	281,54	298,28	316,02	334,81	354,72	375,81	398,16	421,83	446,92	473,49	501,65	27
26	265,58	281,38	298,11	315,83	334,62	354,51	375,59	397,93	421,59	446,66	473,22	501,36	26
25	265,43	281,21	297,94	315,65	334,42	354,31	375,38	397,70	421,35	446,40	472,94	501,07	25
24	265,28	281,05	297,76	315,47	334,23	354,10	375,16	397,47	421,10	446,14	472,67	500,78	24
23	265,12	280,89	297,59	315,29	334,04	353,90	374,94	397,24	420,86	445,88	472,40	500,49	23
22	264,97	280,73	297,42	315,11	333,84	353,69	374,73	397,01	420,62	445,63	472,13	500,20	22
21	264,82	280,57	297,25	314,92	333,65	353,49	374,51	396,78	420,37	445,37	471,85	499,91	21
20	264,67	280,40	297,08	314,74	333,46	353,29	374,29	396,55	420,13	445,11	471,58	499,62	20
19	264,51	280,24	296,91	314,56	333,27	353,08	374,08	396,32	419,89	444,86	471,31	499,33	19
18	264,36	280,08	296,73	314,38	333,07	352,88	373,86	396,09	419,65	444,60	471,04	499,05	18
17	264,21	279,92	296,56	314,20	332,88	352,67	373,65	395,86	419,40	444,34	470,76	498,76	17
16	264,05	279,76	296,39	314,02	332,69	352,47	373,43	395,64	419,16	444,09	470,49	498,47	16
15	263,90	279,59	296,22	313,83	332,50	352,27	373,21	395,41	418,92	443,83	470,22	498,18	15
14	263,75	279,43	296,05	313,65	332,30	352,06	373,00	395,18	418,68	443,57	469,95	497,89	14
13	263,60	279,27	295,88	313,47	332,11	351,86	372,78	394,95	418,43	443,32	469,68	497,61	13
12	263,45	279,11	295,71	313,29	331,92	351,66	372,57	394,72	418,19	443,06	469,41	497,32	12
11	263,29	278,95	295,54	313,11	331,73	351,45	372,35	394,49	417,95	442,80	469,14	497,03	11
10	263,14	278,79	295,37	312,93	331,54	351,25	372,14	394,27	417,71	442,55	468,86	496,74	10
9	262,99	278,63	295,20	312,75	331,35	351,05	371,92	394,04	417,47	442,29	468,59	496,46	9
8	262,84	278,47	295,02	312,57	331,15	350,85	371,71	393,81	417,23	442,04	468,32	496,17	8
7	262,69	278,31	294,85	312,39	330,96	350,64	371,49	393,58	416,99	441,78	468,05	495,88	7
6	262,53	278,14	294,68	312,21	330,77	350,44	371,28	393,36	416,75	441,53	467,78	495,60	6
5	262,38	277,98	294,51	312,03	330,58	350,24	371,06	393,13	416,51	441,27	467,51	495,31	5
4	262,23	277,82	294,34	311,85	330,39	350,04	370,85	392,90	416,27	441,02	467,24	495,03	4
3	262,08	277,66	294,17	311,67	330,20	349,83	370,64	392,68	416,02	440,76	466,97	494,74	3
2	261,93	277,50	294,00	311,49	330,01	349,63	370,42	392,45	415,78	440,51	466,70	494,45	2
1	261,78	277,34	293,83	311,31	329,82	349,43	370,21	392,22	415,54	440,25	466,43	494,17	1
Réf. 0	261,63	277,18	293,66	311,13	329,63	349,23	369,99	392,00	415,30	440,00	466,16	493,88	0
Cent	*do*	*do# réb*	*ré*	*ré# mib*	*mi*	*fa*	*fa# solb*	*sol*	*sol# lab*	*la*	*la# sib*	*si*	Cent

Figure B.3 Fréquences des sons d'une gamme chromatique de l'octave 3. Tempérament égal (la_3 = 440 Hz). Valeurs croissantes, par pas de 1 cent.

ANNEXE C

Pratique du lecteur musicien

1. Fréquence, période, célérité, longueur d'onde

Il est important pour un musicien, peu familier avec le langage physique du son, de s'approprier très rapidement les grandeurs de base et leurs unités : la fréquence ***f***, la période **T**, la célérité du son **c**, la longueur d'onde λ et de connaître les principales relations numériques qui les lient.

f (Hz) **= 1/T** (s) et réciproquement, **T = 1/*f***

λ (m) = **c** (m)/***f*** (Hz) et **λ = cT**

Remarque : la période des sons qui nous concernent en acoustique est toujours inférieure à la seconde. En pratique, on utilise souvent la milliseconde, mais, dans le calcul, il faut impérativement écrire la période en secondes.

2. Correspondances note-fréquence-période-longueur d'onde

Lorsqu'il associe en permanence l'écoute d'un son à sa représentation sonagraphique, un musicien accède rapidement à une première interprétation de l'image. Cependant, pour tirer pleinement parti des analyses et en particulier des données numériques, il devra dès le début :

- développer l'habitude de passer mentalement – et avec agilité – de la fréquence à la note de musique avec son indice dans l'octave,
- passer de la fréquence à la période et à la longueur d'onde ;
- visualiser le tout sur la portée musicale (figure C.1).

Deux repères sont utiles au début : le si_4 (sous le contre-ut) et le sol_1 (première ligne en clé de *fa*).

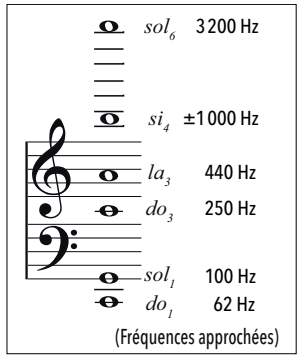


Figure C.1 Correspondance entre notes musicales et fréquences.

La fréquence du si_4 est voisine de 1 000 Hz ; sa période est donc de 1 ms et sa longueur d'onde dans l'air (à 20 °C), de 0,34 m. La fréquence du sol_1 étant voisine de 100 Hz, on en déduit immédiatement : T = 10 ms et λ = 3,40 m. La connaissance de la longueur d'onde permet de se représenter le comportement des ondes sonores en fonction des dimensions des obstacles qu'elles rencontrent.

Pour les autres sons il est aisé de procéder en utilisant l'expression numérique des principaux intervalles, fournie par leur position dans la série harmonique, et en particulier l'octave qui permet de passer rapidement du grave à l'aigu.

3. Série harmonique

C'est un schème fondamental qui régit les intervalles entre les sons et permet de lire rapidement un sonagramme (voir chapitre 1, § 4.4). Il faut connaître par cœur et sans hésitation la série harmonique jusqu'au 16e harmonique (ou plus) en partant d'une note fondamentale quelconque (*sol*, *sib*, *fa*#, *mib*). Mais surtout, écoutez les exemples sonores de cet ouvrage, plusieurs fois, dans de bonnes conditions.

4. Un exemple pratique : calculer la fréquence de résonance d'une bouteille

Figure C.2 Bouteille étoilée certifiée d'une contenance de 1 litre.

Pour entendre la fréquence de résonance d'une bouteille, chacun sait qu'il faut souffler légèrement sur le bord de l'ouverture en recouvrant partiellement l'orifice du goulot avec les lèvres, à la manière d'un joueur de flûte de Pan. Prenons l'exemple d'une bouteille d'un litre (figure C.2). À l'oreille, nous entendons un *sol* grave, ce que confirme la mesure de la fréquence : 104 Hz, soit un sol_2 à 3 cents près. La théorie acoustique du résonateur fournit une relation qui gouverne les principaux paramètres déterminant la fréquence *f*. Ce sont : le volume (V) de la cavité, la longueur (l) du goulot, la section (s) du goulot et la célérité (c) du son. Cette relation est :

$$f = (c/2\pi) \times (s/(V\ l))^{1/2} \quad \text{avec } s = \pi\ (d^2/4)$$

Les mesures de la bouteille sont $l \approx 70$ mm ; $d = 18$ mm. La difficulté consiste à ne pas faire d'erreur au moment de la conversion en mètres.

Sachant que 1 litre = 1 dm^3 on obtient :

$$s = \pi \times 81 \times 10^{-6} \quad avec\ l = 7 \times 10^{-2} \quad et \quad V = 1 \times 10^{-3}$$

Le calcul final pour c = 340 m/s à la température ambiante aboutit à :

$$\boldsymbol{f\ (Hz) = (340/2\pi) \times (8{,}1\pi/7)^{1/2} = 103{,}17\ Hz}$$

Ce résultat correspond pratiquement à la fréquence mesurée acoustiquement. Ce n'est pas toujours le cas. Avec une bouteille d'un demi-litre, l'écart était de 146 Hz (calculé) à 143 Hz (mesuré). La frontière entre le volume principal et le goulot proprement dit est difficile à décider.

On remarquera que, lorsque le volume double, la fréquence est divisée par la racine carrée de 2. Or, racine de 2 est le rapport de fréquence qui existe entre un son et sa quarte augmentée dans le tempérament égal : c'est l'intervalle du triton (voir le tableau des intervalles, page 480).

ANNEXE D

Textes

1. Analyse de l'écoute d'une mélodie par Edmund Husserl

Dans les *Leçons pour une phénoménologie sur la conscience intime du temps* (1905), Husserl analyse le concept de conscience du temps en s'appuyant sur la notion d'objets temporels. Il prend à plusieurs reprises des exemples empruntés à la musique et en particulier celui de l'écoute d'une mélodie[1].

> *Par « objets temporels, au sens spécial du terme », nous entendons des objets qui ne sont pas seulement des unités dans le temps, mais contiennent aussi en eux-mêmes l'extension temporelle. Quand un son résonne, mon appréhension objectivante peut prendre pour objet le son qui dure et résonne là, et non pourtant la durée du son ou le son dans sa durée. Celui-ci, comme tel est un objet temporel. Il en va de même pour une mélodie, pour n'importe quel changement, mais aussi pour toute persistance considérée comme telle. Prenons l'exemple d'une mélodie, ou d'un fragment de mélodie d'un seul tenant. La chose semble tout d'abord fort simple ; nous entendons la mélodie, c'est-à-dire nous la percevons, car entendre, c'est percevoir. Pendant que résonne le premier son, le second arrive, puis le troisième, etc. Ne devons-nous pas dire : quand le second son résonne, alors je l'entends « lui », mais je n'entends plus le premier, etc. ? En vérité je n'entends donc pas la mélodie, mais seulement le son individuel présent. Que le fragment écoulé de la mélodie soit pour moi objectif, j'en suis redevable – ainsi sera-t-on porté à dire – au souvenir ; et que je ne suppose pas, à chaque fois qu'un son se produit, que ce soit là «tout», j'en suis redevable à l'attente pré-voyante. Mais nous ne pouvons nous reposer sur cette explication, car tout ce que nous venons de dire se reporte sur le son individuel. Quand on le fait résonner, je l'entends comme présent, mais pendant qu'il continue à résonner il a un présent toujours nouveau, et le présent chaque fois précédent se change en passé. Je n'entends donc à chaque fois que la phase actuelle du son, et l'objectivité de l'ensemble du son qui dure se constitue dans le « continuum » d'un acte qui, pour une part, est souvenir, pour une part, très petite, ponctuelle, perception, et pour une part plus large, attente.*
>
> Husserl, E., 1964, p. 36-37. © PUF

2. L'illusion perceptive du trille

Le trille, alternance de deux notes, est un ornement instrumental que la voix s'efforce de reproduire, ou du moins d'en donner l'effet perceptif, puisqu'elle ne peut que glisser continument d'une note à l'autre (voir chapitre 6, § 4.4). Stephen de la Madelaine commente avec humour la pratique des artistes de son temps.

---

1. Voir chapitre 4, § 2.3. Le texte original comporte des mots en italique mis entre guillemets ici.

Le trille, qui est pour le chant ce que les diamants sont pour la parure d'une femme, offre avec eux plus d'un rapprochement. Les vrais brillants sont les seuls qui puissent passer dans le commerce ; mais combien n'en voit-on pas qui ursurpent l'admiration du vulgaire ? La fabrication du strass produit des merveilles qui trompent quelquefois l'œil des connaisseurs patentés. Il en est de même du trille qui, exécuté avec aplomb dans une salle de théâtre, où il est difficile d'apprécier nettement les détails d'une vocalisation délicate, prête beaucoup à la contrefaçon et au charlatanisme.

J'ai toujours trouvé qu'une parabole rend la vérité plus saisissante et la démonstration plus facile. J'ai besoin, comme Sancho, de naïve mémoire, qu'on me permette d'exposer les faits à ma manière. Le lecteur y gagnera en lucidité ce qu'il pourrait y perdre d'un temps qui n'est regrettable que quand on le sacrifie sans compensation agréable ou utile.

On m'a parlé d'un saint homme qui avait été en Palestine et qui en avait rapporté une relique précieuse : ce n'était rien moins que l'un des blonds cheveux de la sainte Vierge. Lorsque le digne pèlerin le produisait aux yeux des fidèles, il avait soin de se mettre à une certaine hauteur afin d'être vu de tout le monde ; puis il prenait délicatement son cheveu entre l'index et le pouce de chaque main, et le tendait de toute sa longueur pour montrer sa finesse et son élasticité. Les pieux assistants n'y voyaient pas grand'chose, mais ils se prosternaient devant cette relique invisible avec une confiance qui s'est un peu perdue depuis. Il est vrai de dire que l'homme au cheveu n'y voyait rien de plus que les autres. C'était un chrétien de vieille race, il aurait rougi de commettre une supercherie honteuse en exhibant le cheveu d'une simple pécheresse aux lieu et place de celui qu'il exposait à la vénération de la foule ; et comme en réalité le pèlerin avait perdu l'inestimable relique dans une tempête en mer, il avait trouvé plus simple et beaucoup moins déloyal de se borner à la pantomime qui produisait sur les fidèles agenouillés à distance le même effet que si le prétendu cheveu eût réellement existé.

Je me suis cent fois remémoré cette fraude pieuse en écoutant en fort bons lieux des trilles qui escamotaient si habilement la note supérieure, que le public croyait entendre un pianissimo de la plus exquise délicatesse, tandis qu'en réalité l'artiste faisait sautiller sa voix sur une seule note.

D'autres chanteurs se bornent à soulever la première note du trille, de manière à obtenir une seconde diminuée, d'une justesse équivoque, et le rapide martellement du trait (quand il y a martellement) produit encore, dans cette circonstance, une illusion qui satisfait beaucoup d'oreilles. Mais le seul trille régulier, tel que le battent madame Damoreau et Ponchard, est celui qui produit nettement et vigoureusement les deux notes qui le composent.

Madelaine, (de la) S., vers 1851, p. 119-121.

ANNEXE E

Visualisation des phénomènes vibratoires

Les mouvements vibratoires qui sont à la base de la production des sons sont d'amplitude très faible et généralement beaucoup trop rapides pour être visibles à l'oeil nu. Nous disposons aujourd'hui de plusieurs techniques pour accéder visuellement au détail des mouvements. Les deux principales sont d'une part le ralentissement des images captant le mouvement réel, voire le rendant immobile par stroboscopie, et d'autre part les simulations animées rendues accessibles grace aux possibilités de l'informatique. Les simulations permettent d'augmenter considérablement l'amplitude des mouvements et, pour certaines, de manipuler soi-même les paramètres de la vibration pour voir et entendre leur incidence sur le comportement global. Parmi les sites disponibles nous en avons sélectionné quatre dont la présence en ligne est stable depuis quelques années.

1. À propos des animations visibles en ligne

1.1. *sDR* : site de Dan Russel[1]

<http://www.gmi.edu/~drussell/Demos.html>

Ce site offre de nombreuses animations démontrant les données fondamentales de l'acoustique et des vibrations. Voir en particulier :

- Rayonnement (*radiation*) :
  http://www.acs.psu.edu/drussell/Demos/rad2/mdq.html
- Modes vibratoires d'un diapason (*tuning fork*) :
  http://www.acs.psu.edu/drussell/Demos/TuningFork/fork-modes.html
- Modes vibratoires de guitares :
  http://www.acs.psu.edu/drussell/guitars/index.html
- Mouvement d'une corde pincée (*Motion of a Plucked String*) par capture d'images grande vitesse (1 000 images par seconde) permettant de voir le mouvement de la corde au ralenti (27''), un aller et retour en boucle (1'04) puis le mouvement global très ralenti (1'10).

Démonstrations en ligne.

1. Penn State, États-Unis.

1.2. *sPF* : site de Paul Falstad

<http://www.falstad.com/mathphysics.html>

Ce site personnel offre d'ingénieuses applications (*applets*) qui ont la particularité d'être interactives : l'utilisateur peut modifier les paramètres qui régissent le mouvement. Voici un exemple (figure E.1).

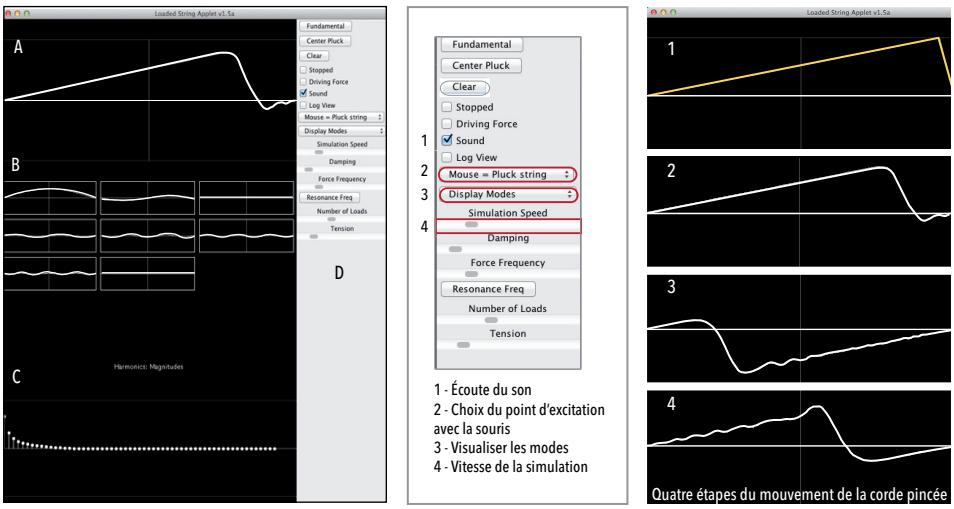


Figure E.1 Animation du mouvement d'une corde pincée. À gauche : zones d'une fenêtre de l'applet. De haut en bas, mouvement de la corde (A), mouvements des modes propres séparés (B), équivalent de l'amplitude des harmoniques (C), barre de réglage des paramètres (D). Au centre : réglages à effectuer pour la démonstration décrite. À droite : aspects de la corde au cours du mouvement animé.

Sélectionnez une animation dans la liste, par exemple *Loaded String applet*, et acceptez Java. L'animation apparaît dans une nouvelle fenêtre ; pour stopper le mouvement, cliquez sur *Clear* dans la barre de commande, puis modifiez les réglages selon les indications de la figure E.1 (partie centrale) en veillant bien à ralentir la vitesse d'animation. Vous pouvez alors cliquer sur la corde et lui donner une forme triangulaire en déplaçant la souris vers le haut, près d'une extrémité, et voir la forme que prend la corde pendant la propagation du mouvement. En choisissant *Mouse = Shape string*, vous pourrez donner à la corde une forme vibratoire quelconque. Les actions sont similaires avec les applets *Rectangular membrane* et *Circular membrane*.

Il est préférable de télécharger l'applet pour l'installer sur l'ordinateur. Ouvrir ensuite le fichier « index.htm ».

1.3. *sJW* : site de Joe Wolfe[2]

<http://www.animations.physics.unsw.edu.au/waves-sound/>

Ce site très complet et pédagogique est un cours de physique des phénomènes ondulatoires. Il offre aussi quelques démonstrations sur le fonctionnement des instruments de musique. Toutes les animations sont commentées en anglais.

À l'ouverture de la page *Waves & sound*, choisir une animation, par exemple *The Doppler effect,* puis attendre le téléchargement de la page, qui est de l'ordre de la

2. Université de Sydney, Australie.

minute (voir la barre de progression). Ensuite sélectionner un chapitre dans la fenêtre de droite.

Un grand nombre de démonstrations sont téléchargeables selon deux formats : .swf (à lire avec Flash Player) et .html (à déposer sur la page blanche d'un navigateur).

1.4. *sWR* : site de Wolfgang et Rousseau[3]

<*http://webapps.fundp.ac.be/didactique/ondes/index.php*>

Ce site propose quelques animations démontrant les phénomènes fondamentaux des ondes en lien avec les données physiques : onde transversale, longitudinale, addition d'ondes, réflection, etc. Démonstrations en ligne.

2. Fibrostroboscopie des cordes vocales pendant le chant

La fibroscopie permet d'accéder au fond du pharynx pour éclairer et filmer les mouvements des cordes vocales. On peut y associer la stroboscopie obtenue grâce un éclairage intermittent de flashes lumineux que l'on peut synchroniser à la fréquence vibratoire du système – ce qui produit une image immobile – ou à une fréquence voisine, ce qui produit un mouvement ralenti de l'observation. La stroboscopie requiert un son tenu de fréquence stable.

Nous avons sélectionné quatre documents :

- <https://www.youtube.com/watch?v=v9Wdf-RwLcs>
  (Université de Washington, durée 2'20)
  Démonstration d'une stroboscopie (voix féminine), avec diverses hauteurs de son. Au début *sol*#$_3$; vers 1'07 note grave (si_2) ; vers 1'17 note aiguë (si_3); puis sons plus ou moins forts.
- <https://www.youtube.com/watch?v=9Tlpkdq8a8c>
  (Dr J.-P. Thomas, durée 1 minute)
  Stroboscopie des cordes vocales (voix féminine) lors de la production de sons graves, aigus et au cours d'un glissando
- <https://www.youtube.com/watch?v=mJedwz_r2Pc>
  Passer directement à la section 0'20 à 0'50 qui montre un très beau ralentissement mettant en évidence le mouvement d'ouverture des cordes vocales dans leur épaisseur : du bord inférieur au bord supérieur.
- <https://www.youtube.com/watch?v=-XGds2GAvGQ>
  (Sara Lundberg, durée 3 minutes)
  Cet étrange film, qui est une fibroscopie simple (sans stroboscopie), donne à voir l'image captée par la caméra miniature située à l'extrémité de la fibre au cours de l'interprétation d'une pièce polyphonique de Luis de Victoria par un quatuor vocal. Il commence par l'introduction de la fibre dans une des narines de chaque chanteur, puis l'écran présente les quatre larynx situés côte à côte pendant l'interprétation de la pièce. Au début, les cordes vocales de chaque chanteur sont écartées. On peut ensuite suivre les entrées de chacun d'eux au cours de l'exécution en repérant la fermeture des cordes vocales de celui ou de celle qui chante.

3. Université de Namur.

ANNEXE F

Bibliographie

1. Matériel audiovisuel

Arom, S., 1970, *L'arc musical ngbaka* (film 16 mm, numérisé). République Centrafricaine, CNRS (consultable sur : http://videotheque.cnrs.fr/doc=457).

Asselin, P.-Y., & Legaillard, Y., 1985, *Musique et tempérament. Accord de l'orgue et du clavecin selon les tempéraments anciens* (2 musicassettes). Paris, ERATO/CNRS/Costallat.

Auteurs divers, 1996, *Les Voix du monde ; une anthologie des expressions vocales.* Collection CNRS/Musée de l'Homme (3 CD + livret édités par H. Zemp, G. Léothaud, B. Lortat-Jacob), Paris, Le Chant du Monde (http://archives.crem-cnrs.fr).

Bregman, A. S., & Ahad, P. A., 1996, CD, *Demonstrations of Auditory Scene Analysis: The Perceptual Organization of sound*. Montréal (Canada), distribué par MIT Press.

Deutsch, D., 1995, CD, *Musical Illusions and Paradoxes*. Philomel.com.

Duda, R. O., 1996, exemples sonores accompagnant l'article : « Auditory localization demonstrations ». *Acustica*, vol. 82, p. 346-355 (CD de 48 exemples).

Houtsma, A., Rossing, T. D., & Wagenaars, W. M., 1987, CD, *Auditory demonstrations*. Eindhoven (Pays-Bas), distribué par Acoustical Society of America.

Kaltenecker, M., 1993, notice du CD *Helmut Lachenmann* (Bohy, F., trad.). Una Corda, CD 202082.

Risset, J.-C., 1995, *An introductory catalog of computer-synthesized sounds*. CD (Réédition du texte et des exemples sonores, 1969). Computer Music Currents 13, WER 20332. Mayence (Allemagne), Wergo.

Schaeffer, P., & Reibel, G., 1967, *Solfège de l'objet sonore* (disques 33 t. et livret), réédition 1998 (3 CD). Paris, Ina/GRM.

Suva PRO., 1997, CD *Audio Demo 3*. Lucerne (Suisse).

2. Bases de données

Iowa, Fritts L. & coll. http://theremin.music.uiowa.edu/MIS.html
Musical instruments samples (MIS), 1997-2011 mono. Depuis 2013 en évolution. Téléchargement libre.

Ircam, Base de données de sons musicaux « SOL », 2006, intégrée à ORCHIDS : http://forum-net.ircam.fr/fr/product/orchids-2/
Nombreux modes de jeu utilisés en musique contemporaine.
© Ircam-Centre Pompidou.

McGill, McGill University Master Samples (MUMS), 11 CD (ne sont plus disponibles)
© Franck Opolko et Joel Wapnick (1987-1988).

RWC, Masataka Goto et coll. 2001, https://staff.aist.go.jp/m.goto/RWC-MDB/rwc-mdb-i.html
RWC-Music Database: Musical Instrument Sound : 50 DVD, plusieurs interprètes, plusieurs modèles d'un même instrument, plusieurs nuances. Instruments de la musique traditionnelle japonaise. © 1994-2015 by Masataka Goto.

3. Livres – Périodiques – Articles – Comptes rendus de congrès – Thèses – Rapports

AFNOR, 1972, « Vocabulaire de l'acoustique musicale » (vol. Norme NF S30-107). Paris, AFNOR (Association française de normalisation).

Andler, D., 2002, « La forme ». In Andler, D., Fagot-Largeault, A., & Saint-Sernin, B. (éd.), *Philosophie des Sciences*. Paris, Gallimard, coll. « Folio-Essais », vol. 2, p. 1049-1130.

ANSI, 1973, « Psychoacoustical terminology » (vol. S3.20-1973, Washington, ANSI (American National Standard Institute).

Aran, J. M., 1988, « Préface ». In Collectif, *Physiologie de la cochlée*. Paris, INSERM/SFA/CNET.

Arnaut de Zwolle, H., vers 1450, *Les traités d'Henri-Arnaut de Zwolle et de divers anonymes*. Ms B.N. Latin 7295. Kassel, Bärenreiter Verlag, fac-similé 1952.

Arom, S., Fernando, N., Fürniss, S., Le Bonin, S., *et al.*, 2007, « Typologie des techniques polyphoniques ». In Nattiez, J. J. (éd.), *Musiques – Une encyclopédie pour le XXIe siècle*. Paris, Actes Sud-Cité de la Musique, vol. 5 *L'unité de la musique*, p. 1088-1109.

Arom, S., Léothaud, G., & Voisin, F., 1997, « Experimental ethnomusicology. An interactiv approach to the study of musical scales ». In Deliège, I., & Sloboda, J. (éd.), *Perception and Cognition of Music*. Hove (Royaume-Uni), Erlbaum (Taylor and Francis Group), p. 3-30.

Artaud, P.-Y., & Geay, G., 1980, *Flûtes au présent – Present Day Flutes. Traité des techniques contemporaines sur les flûtes traversières à l'usage des compositeurs et des flûtistes* (avec disque). Paris, Jobert & Éd. Musicales Transatlantiques (1re éd.), G. Billaudot (2e éd.), 133 p.

Assayag, G., & Cholleton, J. P., 1995, « Musique, nombres et ordinateurs ». *La Recherche,* 278, p. 804-809.

Assayag, G., Castellengo, M., & Malherbe, C., 1985, *Nouvelles techniques instrumentales : composition et formalisation*, Rapport de recherche n° 38, Paris, IRCAM.

Asselin, P.-Y., 1983, *Compréhension et réalisation des systèmes d'accord à l'ancienne*. Thèse de l'université Paris 6.

Asselin, P.-Y., 1985, *Musique et tempérament – Théorie et pratique de l'accord à l'ancienne*. Paris, Éditions Costallat (1re éd. avec 2 cassettes audio), 236 p.

Bachem, A., 1937, « Various types of absolute pitch ». *J. Acoust. Soc. Am.,* 11, p. 434-439.

Baddeley, A. D., 1993, *La mémoire humaine. Théorie et pratique*. Presses universitaires de Grenoble, 547 p.

Bailblé, C., 1999, *La perception et l'attention modifiées par le dispositif cinéma*. Thèse de l'université Paris 8.

Bailly, L., Henrich, N., Pelorson, X., & Gilbert, J., 2008, *Vocal folds and ventricular bands in interaction: comparison between "in vivo" measurements and theoretical predictions*. Conférence présentée à Acoustics'08, Paris.

Ballas, J. A., 1999, « The interpretation of natural sounds in the cockpit ». In Stanton, N., & Edworthy, J. (éd.), *Human factors in auditory warnings*. Ashgate, p. 91-112.

Baltrusaitis, J., 1969, *Anamorphoses, ou magie artificielle des effets merveilleux*. Paris, Olivier Perrin, 186 p.

Barbaras, R., 1994, *La perception, essai sur le sensible*. Paris, Hatier, 79 p.

Barbour, J. M., 1972, *Tuning and Temperament – A Historical Survey*. New York, Da Capo Press Inc., 228 p.

Barthélemy, J.-P., & Guénoche, A., 1988, *Les arbres et les représentations des proximités*. Paris, Masson, 240 p.

Barthélemy, J.-P., 1993, « Similitude, arbres et typicalité ». In Dubois, D. (éd.), *Sémantique et cognition – Catégories, prototypes, typicalité*. Paris, CNRS, p. 205-224.

Bartolozzi, B., 1967, *New sounds for woodwind* (+ disque). Londres, Oxford University Press, 78 p.

Beauchamp, J. W., & Lakatos, S., 2002, *New spectro-temporal measures of musical instrument sounds used for a study of timbral similarity of rise-time and centroid-normalized musical sounds*. Conférence présentée à la 7th ICMPC (International Conference on Music Perception and Cognition), Sydney.

Bédos de Celles (Dom), F., 1766, *L'art du facteur d'orgues* (vol. 1). Fac-similé, 1963, Kassel, Bärenreiter Verlag, 142 p.

Bédos de Celles (Dom), F., 1770, *L'art du facteur d'orgues* (vol. 2/3). Fac-similé, 1965, Kassel, Bärenreiter Verlag, 393 p.

Behnke, E., 1890 (9e éd.), *The mechanism of the human voice*. Londres, Curwen & sons, 156 p.

Békésy (von), G., 1960, *Experiments in Hearing* (Wever, E. G., trad.). New York, McGraw-Hill Book Company Inc., 745 p.

Bennett, G., 1992, « The sense of the Phrase – Compositional Grouping in Music » (avec CD). In Sundberg, J. (éd.), *Gluing tones*. Stockholm, Royal Swedish Academy of Music, p. 9-25.

Bensa, J., Dubois, D., & Kronland-Martinet, R., 2005, « Perception and cognitive evaluation of a piano synthesis model ». *Lecture Notes in Computer Science,* 3310, p. 232-245.

Berger, K. W., 1964, « Some factors in the recognition of timbre ». *J. Acoust. Soc. Am.,* 36 (10), p. 1888-1891.

Berlioz, H., 1855 (2e éd.), *Traité d'instrumentation et d'orchestration – L'art du chef d'orchestre*. Fac-similé s.d., Paris, Éditions Henry Lemoine, 312 p.

Berthoz, A., & Petit, J.-L., 2006, *Phénoménologie et physiologie de l'action*. Paris, Odile Jacob, 350 p.

Berthoz, A., 1997, *Le sens du mouvement*. Paris, Odile Jacob, 205 p.

Bertrand, D., 1992, *Les chevalets plats de la lutherie de l'Inde* (vol. 9). Paris, Maison des sciences de l'homme, 96 p.

Beyhom, A., 2010, *Théories de l'échelle et pratiques mélodiques chez les Arabes. Une approche systématique et diachronique* (vol. 1). Paris, Geuthner, 674 p.

Bideaud, J., & Houdé, O., 1989, « Le développement des catégorisations : capture logique ou capture écologique des propriétés des objets ». *L'Année psychologique,* 89, p. 87-123.

Biot, J. M., 1817, *Précis élémentaire de physique expérimentale*, vol. III *De l'acoustique*. Paris, Deterville.

Bismarck (von), G., 1974, « Sharpness as an attribute of the timbre of steady sounds ». *Acustica,* 30, p. 159-172.

Bismarck (von), G., 1974, « Timbre of steady sounds: A factorial investigation of its verbal attributes ». *Acustica,* 30, p. 146-159.

Blacking, J., 1973, *Le sens musical*. Paris, Éditions de Minuit, 130 p.

Blades, J., 1961 (2e éd.), *Orchestral Percussion Technique*. Londres, Oxford University Press, 85 p.

Blauert, J., 1996, *Spatial Hearing – The Psychophysics of Human Sound Localization* (Allen, J. S., trad. 2e éd.). Cambridge, The MIT Press, 494 p.

Bonnet, C., 1986, *Manuel pratique de psychophysique*. Paris, Armand Colin, 254 p.

Borel-Maisonny, S., 1976, « Étude radiographique des mouvements oro-pharyngés pendant la parole et le jeu instrumental ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 86. LAM, université Paris 6.

Boring, E. G., 1930, « A New Ambiguous Figure ». *American Journal of Psychology,* 42, p. 444.

Bosseur, J.-Y., 2005, *Du son au signe – Histoire de la notation musicale*. Paris, Éditions Alternatives, 143 p.

Botte, M. C., 1989, « Perception de l'intensité sonore ». In Collectif, *Psychoacoustique et perception auditive*. Paris, Inserm/SFA/CNET, p. 13-41.

Bouasse, H., 1906, *Bases physiques de la musique* (vol. 28, *Scientia*). Paris, Gauthier-Villars, 112 p.

Bouasse, H., 1926, *Acoustique générale – Ondes aériennes*. Paris, Librairie Delagrave (fac-similé, Paris, Blanchard, 1987), 544 p.

Bouasse, H., 1926, *Cordes et membranes*. Paris, Librairie Delagrave, 505 p.

Bouasse, H., 1927, *Verges et plaques, cloches et carillons*. Paris, Librairie Delagrave, 455 p.

Bouasse, H., 1929, *Instruments à vent* (vol. 1). Paris, Librairie Delagrave, 410 p.

Bouasse, H., 1929, *Tuyaux et résonateurs – Introduction à l'étude des instruments à vent*. Paris, Librairie Delagrave (fac-similé, Paris, Blanchard, 1987), 491 p.

Bouasse, H., 1986, *Instruments à vent* (vol. 2). Paris, Librairie A. Blanchard, 386 p.

Bougeret, G., 1982, *Échelle des hauteurs et musique instrumentale à la Renaissance.* Doctorat de 3e cycle. Université François-Rabelais, Tours.

Boulez, P., 1991, « Le timbre et l'écriture, le timbre et le langage ». In Barrière, J.-B. (éd.), *Le timbre, métaphore pour la composition*. Paris, Christian Bourgois/IRCAM, p. 541-549.

Boulez, P., 1984, « Un entretien avec Pierre Boulez », in *Le Monde* daté du 14 novembre, interview de Claude Samuel, cahier spécial « Forum des percussions », 14 novembre 1984-6 janvier 1985.

Boulez, P., & Cage, J., 2002, *Correspondance et Documents.* Édités par Jean-Jacques Nattiez, nouvelle édition revue par Robert Piencikowski. Mayence (Allemagne), Schott, 360 p.

Bourcet, P., & Lienard, P., 2002, « Acoustique fondamentale ». In Mercier, D. (éd.), *Le livre des techniques du son*, vol. 1 (3e éd.). Paris, Dunod, p. 1-43.

Brailoiu, C., 1953, « Sur une mélodie russe ». In *Musique russe*, vol. 2. Paris, PUF, p. 329-391.

Bregman, A. S., 1990, *Auditory Scene Analysis : The Perceptual Organization of Sound*. Cambridge, Massachusetts, MIT Press, 773 p.

Browne, L., & Behnke, E., 1983, *La voix, le chant et la parole* (Garnault, D., trad.). Paris, Société d'édition scientifique, 328 p.

Bruneau, M., 1998, *Manuel d'acoustique fondamentale*. Paris, Hermès, 576 p.

Burns, E. M., 1999, « Intervals, scales, and tuning ». In Deutsch, D. (éd.), *The psychology of music* (2e éd.). New York, Academic Press, p. 215-264.

Buser, P., & Lestienne, R., 2001, *Cerveau, Information, Connaissance*. Paris, CNRS Éditions, 228 p.

Cadoz, C., 1991, « Timbre et causalité ». In Barrière, J.-B. (éd.), *Le timbre, métaphore pour la composition*. Paris, Christian Bourgois/IRCAM, p. 17-46.

Cage, J., 1994, *Je n'ai jamais écouté aucun son sans l'aimer : le seul problème avec les sons, c'est la musique*. (Charles, D., trad.). La Souterraine (France), La main courante, 30 p.

Callas, M.-F., & Fontaine, J.-M., 1996, *La conservation des documents sonores*. Paris, CNRS, 208 p.

Canévet, G., 1989, « Audition binaurale et localisation auditive : aspects physiques et psychoacoustiques ». In Collectif, *Psychoacoustique et perception auditive*. Paris, Inserm/SFA/CNET, p. 83-122.

Carpentier, G., 2008, *Approche computationnelle de l'orchestration musicale*. Thèse de l'université Paris 6.

Castanet, P.-A., 2007, « Edgard Varèse et la "philosophie du bruit" – Ionisation en questions ». In Horodyski, T., & Lalitte, P. (éd.), *Edgard Varèse – Du son organisé aux arts audio*. Paris, L'Harmattan, p. 39-70.

Castellengo, M., 1993, *Fusion or separation: from vibrato to vocal trill.* Conférence présentée au SMAC93 (Stockholm Music Acoustics Conference), Stockholm.

Castellengo, M., 1994, « La perception auditive des sons musicaux ». In Zenatti, A. (éd.), *Psychologie de la musique.* Paris, PUF, p. 55-86.

Castellengo, M., 1994, « Les formes sonores ». In Noël, É. (éd.), *Les sciences de la forme aujourd'hui*. Paris, Seuil, Point sciences, p. 126-139.

Castellengo, M., 1999, « Analysis of initial transients in flute-like instruments ». *Acta Acustica & Acustica,* 85, p. 387-400.

Castellengo, M., 2004, *Flute-like instrument transients: An analytical study of mouth-tone production versus pressure rise time.* Conférence présentée à l'ISMA (International Symposium on Musical Acoustics), Nara, Japon.

Castellengo, M., 2010, « Les sources acoustiques ». In Mercier, D. (éd.), *Le livre des techniques du son*, vol. 1 (4e éd.). Paris, Dunod, p. 45-80.

Castellengo, M., 2014, « Perception(s) de la voix chantée : une introduction ». In Henrich-Bernardoni, N. (éd.), *La voix chantée entre sciences et pratiques*. Paris, Bruxelles, De Boeck-Solal, p. 35-64.

Castellengo, M., D'Alessandro, C., & Richard, G., 1989, *Study of vocal pitch vibrato perception using synthesis.* Conférence présentée à l'ICA (13th International Congress on Acoustics), Belgrade.

Castellengo, M., Drouin, F., & Séchet, P., 1978, « La flûte traversière à une clé, dite flûte baroque ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 97. LAM, université Paris 6.

Castellengo, M., & Dubois, D., 2007, « Timbre ou timbres ? Propriété du signal, de l'instrument ou construction(s) cognitive(s) ? » *Les Cahiers de la SQRM (Société québécoise de recherche en musique),* 9 (1-2), p. 25-38.

Castellengo, M., & Fabre, B., 1994, « The contemporary transverse flute and the shakuhachi : convergences. An acoustic analysis of performance techniques ». *Contemporary Music Review,* 8 (2) : p. 217-237.

Castellengo, M., & Fabre, M., 1966, « Le galoubet et le tambourin de Provence ». *Bulletin du GAM (Groupe d'acoustique musicale),* n°23. LAM, université Paris 6.

Castellengo, M., & Forest, L., 1999, « Métamorphoses de la flûte traversière au XIXe siècle : esthétique musicale, acoustique et facture ». In *Acoustique et instruments anciens, factures, musiques et sciences* (livre + CD). Paris, SFA, Cité de la Musique, p. 85-102.

Castellengo, M., & Goad, P. J., 1996. *Rapport d'expertise sur l'orgue de la cathédrale de Dijon. Étude acoustique de la qualité sonore de jeux d'anches, par comparaison avec les jeux correspondants des orgues de Dole, Poitiers et Saint-Maximin*, Paris, Ministère de la Culture, Direction du Patrimoine.

Castellengo, M., Guyot, F., & Viollon, S., 1996, *Perceptive characterisation of the acoustical quality of real complex sounds – Validation with synthesis*. Conférence présentée au Forum Acousticum EAA, Anvers.

Castellengo, M., Lortat-Jacob, B., & Léothaud, G., 2001, *Pitch perception : five voices with four sardinian singers.* Conférence présentée à l'ISMA (International Symposium on Musical Acoustics), Pérouse.

Caus (de), S., 1615, *Institution harmonique divisée en deux parties : en la premiere sont monstrées les proportions des intervalles harmoniques, et en la deuxiesme les compositions d'icelles*. Francfort, Jan Norton.

Chaigne, A., & Kergomard, J., 2008, *Acoustique des instruments de musique*. Paris, Belin, 704 p.

Chailley, J., 1950, *Histoire musicale du Moyen Âge*. Paris, PUF, 356 p.

Chailley, J., 1954-1955, *Formation et transformation du langage musical*, vol. I. *Intervalles et échelles*. Paris, CDU, Sorbonne, 212 p.

Chailley, J., 1985, *Éléments de philologie musicale*. Paris, A. Leduc, 180 p.

Charbon, P., 1981, *La machine parlante*. Strasbourg, Éditions Jean-Pierre Gyss, 207 p.

Charbonnier, G., & Halbreich, H., 1970, *Entretiens avec Edgard Varèse,* suivis d'une étude de l'œuvre par Harry Halbreich. Paris, Belfond, 171 p.

Cheminée, P., 2009, « Est-ce bien "clair"? Stabilité, instabilité et polysémie d'une forme lexicale en contexte ». In Dubois, D. (éd.), *Le sentir et le dire.* Paris, L'Harmattan, p. 309-338.

Cheveigné (de), A., 2005, « Pitch models ». In Plack, C., & Oxenham, A. (éd.), *Pitch*. New York, Springer.

Chi, T., Ru, P., & Shamma, S. A., 2005, « Multiresolution spectrotemporal analysis of complex sounds ». *J. Acoust. Soc. Am.,* 118 (2), p. 887-906.

Chion, M., 1986, *Le son au cinéma*. Paris, Cahiers du cinéma/Éditions de l'Étoile, 220 p.

Chion, M., 2010 (2e éd.), *Le son – Traité d'acoulogie*. Paris, Armand Colin, 272 p.

Chladni, E., 1809, *Traité d'acoustique* (avec huit planches). Paris, Chez Courcier imprimeur-libraire pour les mathématiques.

Chowning, J., 1997, « The synthesis of complex audio spectra by means of frequency modulation ». *Computer Music Journal,* 1 (2), p. 46-54.

Clark, M., & Luce, D., 1965, « Intensities of Orchestral Instrument Scales Played at Prescribed Dynamic Markings ». *J. Audio Eng. Soc.* (13), p. 151-157.

Clark, M., & Milner, P., 1964, « Dependence of timbre on the tonal loudness produced by musical instruments ». *J. Audio Eng. Soc.,* 12 (1), p. 28-31.

Clark, M., Robertson, P., & Luce, D., 1964, « A preliminary experiment on the perceptual basis for musical instrument families ». *J. Audio Eng. Soc.,* 12 (3), p. 199-203.

Cogan, R., 1969, « Toward a theory of timbre: verbal timbre and musical line in Purcell, Session, and Stravinsky ». *Perspectives of New Music,* VIII, p. 75-89.

Cogan, R., 1984, *New images of musical sound*. Cambridge, Harvard University Press, 177 p.

Cohen, E., 2011. *Le rôle de l'articulation dans l'interprétation de la musique de style baroque pour les instruments à clavier sans dynamique.* Mémoire de DFS, CNSMDP, Paris.

Collectif – Barrière, J.-B. (éd.), 1991, *Le timbre – Métaphore pour la composition*. Paris, Christian Bourgois/IRCAM.

Collectif – Beauchamp, J. W. (éd.), 2007, *Analysis, Synthesis, and Perception of Musical Sounds – The sound of music*. New York, Springer.

Collectif – Botte, M. C., Canévet, G., Demany, L., Sorin, C. (éd.), 1989, *Psychoacoustique et perception auditive*. Paris, Inserm/SFA/CNET/Éditions médicales internationales.

Collectif – Chailley, J. (coord.), 1963, *La résonance dans les échelles musicales.* Conférences présentées au colloque international du CNRS, 9-14 mai 1960 (vol. + 2 disques 33 t). Paris, CNRS, 400 p.

Collectif – Cornut, G. (coord.), 2002, *Moyens d'investigation et pédagogie de la voix chantée*. Lyon, Symétrie, 192 p. + CD-Rom.

Collectif – Dallet, S., & Veitl, A. (éd.), 2001. *Du sonore au musical, Cinquante années de recherches concrètes (1948-1998).* Paris, L'Harmattan, 462 p.

Collectif – Dejonckere, P. H. (éd.), 1995, *Vibrato*. San Diego, Singular Publishing Group Inc.

Collectif – Deliège, I., & Sloboda, J. (éd.), 1994, *Perception and Cognition of Music*. New York, Psychology Press.

Collectif – Deutsch, D., (éd.), 1982, *The Psychology of Music.* San Diego, Academic Press, 542 p.

Collectif – Deutsch, D., (éd.), 1999, *The Psychology of Music* (2e édition). San Diego, Academic Press, 808 p.

Collectif – Dubois, D. (éd.), 2009, *Le sentir et le dire – Concepts et méthodes en psychologie et linguistique cognitives*. Paris, L'Harmattan.

Collectif – Hutchins, C. M. (éd.), 1980, *Sons et musique*. Paris, Belin.

Collectif – Le Ny, J.-F., & Gineste, M.-D. (éd.), 1995, *La psychologie*. Paris, Larousse.

Collectif – McAdams, S., & Bigand, E. (éd.), 1994, *Penser les sons – Psychologie cognitive de l'audition*. Paris, PUF, 402 p.

Collectif – Rossing, T. D. (éd.), 1984, *Acoustics of Bells*. New York, Van Nostrand Reinhold Company.

Collectif – Sundberg, J. (éd.), 1992, *Gluing tones – Grouping in music composition, performance and listening* (+ CD). Stockholm, Royal Swedish Academy of Music.

Collectif, 1694, *Dictionnaire de l'Académie française*. Paris, Jean-Baptiste Coignard.

Collectif, 1988, *Actes des Journées pédagogiques d'acoustique musicale*. Paris, CENAM (disponible au LAM-Paris), 112 p.

Collectif, 1998, *Actes du colloque Acoustique et instruments anciens, factures, musiques, sciences*. Paris, SFA/Musée de la Musique, 270 p. + CD.

Collectif, 2000, « L'Afrique et l'Europe médiévale : la théorie du pentatonisme revue à travers les systèmes africains de tradition orale ». *Musicae Scientiae* (Numéro spécial, Forum de discussion 1), 135 p.

Collectif-ITEMM, 2004, « Qualité sonore des instruments de musique : premières réflexions ». *Musique et technique,* n° 0, p. 15-30.

Collectif-ITEMM, 2007, « Colloque sur la qualité des instruments de musique ». *Musique et technique,* n° 2, p. 12-54.

Collectif-MIM, 1996, *Les unités sémiotiques temporelles, éléments nouveaux d'analyse musicale* (livre + CD). Marseille, Laboratoire MIM, 96 p.

Colomb, C., 1878, *La musique*. Paris, Librairie Hachette, 350 p.

Cornut, G., 2002, « La vibration laryngée ». In Cornut, G. (éd.), *Moyens d'investigation et pédagogie de la voix chantée*. Lyon, Symétrie, p. 9-17.

Cornut, G., 2009 (8e éd.), *La voix*. Paris, PUF, coll. Que sais-je ?, n° 627, 128 p.

Couprie, P., 2001, « Le vocabulaire de l'objet sonore ». In Dallet, S., & Veitl, A. (éd.), *Du sonore au musical. Cinquante années de recherches concrètes (1948-1998).* Paris, L'Harmattan, p. 203-225.

Culin, O., 2006, *L'image musique*. Paris, Fayard, 168 p.

Cuvillier, A., 1956, *Vocabulaire philosophique*. Paris, Librairie Armand Colin, 256 p.

Cyrulnik, B., 1995, *La naissance du sens*. Paris, Hachette, 170 p.

D'Alessandro, C., & Castellengo, M., 1993, *The pitch of short-duration vibrato tones: experimental data and numerical model.* Conférence présentée au SMAC93 (Stockholm Music Acoustics Conference), Stockholm.

D'Alessandro, C., & Castellengo, M., 1994, « The pitch of short-duration vibrato tones ». *J. Acoust. Soc. Am.,* 95 (3), p. 1617-1630.

Dampierre (de), E., 1998, *L'humanité des harpes*. Nanterre, Société d'ethnologie, 47 p.

Damsté, P., 1966, « Vocal cord vibrations compared to the vibrations of a trombone player's lips » (avec film vidéo). *Journal français d'oto-rhino-laryngologie et chirurgie,* 15 (4), p. 395-396.

Dargie, D., 1993, *Thembu Xhosa umngqokolo overtone singing: the use of the human voice as a type of "musical bow".* Conférence présentée à l'International Concil for Traditional Music (ICTM), Berlin.

David, S., 1997, « Représentation d'objets sensoriels et marques de la personne ». In Dubois, D. (éd.), *Catégorisation et cognition : contrastes entre olfaction et audition*. Paris, Kimé, p. 211-242.

Delalande, F., 1996, « Les unités sémiotiques temporelles : problématique et essai de définition ». In *Les unités sémiotiques temporelles, éléments nouveaux d'analyse musicale* (livre + CD). Marseille, Laboratoire MIM, p. 17-25.

Delalande, F., 2001, *Le SON des musiques. Entre technologies et esthétique*. Paris, Ina-GRM/ Buchet-Chastel, 196 p.

Delorme, A., & Flückiger, M. (éd.), 2003, *Perception et réalité. Une introduction à la psychologie des perceptions*. Bruxelles, De Boeck.

Delorme, A., 1982, *Psychologie de la perception*. Paris, Vigot, 421 p.

Demany, L., 1989, « Perception de la hauteur tonale ». In Botte, M. C. (éd.), *Psychoacoustique et perception auditive*. Paris, Inserm/SFA/CNET, p. 43-81.

Descartes, R., 1618. *Abrégé de musique* suivi des *Éclaircissements physiques sur la musique de Descartes du R.P. Nicolas Poisson* (Dumont, P., trad.). Paris, Méridiens Klincksieck, 1990.

Deshays, D., 2006, *Pour une écriture du son*. Paris, Klincksieck, 190 p.

Deutsch, D., 1994, « La perception des structures musicales ». In Zenatti, A. (éd.), *Psychologie de la musique.* Paris, Seuil, p. 115-144.

Deutsch, D., 1999, « Grouping mechanisms in music ». In Deutsch, D. (éd.), *The psychology of Music.* New York, Academic Press Inc., San Diego, p. 299-343.

Diderot, D., & d'Alembert, J. (éd.), 1751-1772, *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers*. Paris, Le Breton, Durand, Briasson, Michel-Antoine David.

Donnadieu, S., 1997, *Représentation mentale du timbre des sons complexes et effets de contexte*. Thèse de l'université Paris 5.

Donnadieu, S., 2007, « Mental representation of the timbre of complex sounds ». In Beauchamp, J. (éd.), *Analysis, Synthesis, and Perception of Musical Sounds – The sound of music*. New York, Springer, p. 272-313.

Doval, B., 2004, « Méthodes d'analyse du signal musical ». In Pachet, F., & Briot, J.-P. (éd.), *Informatique musicale : du signal au signe musical*. Paris, Hermès, Lavoisier, p. 21-65.

Dubois, D., 1993, « Catégorisation et cognition : "10 ans après", une évaluation des concepts de Rosch ». In Dubois, D. (éd.), *Sémantique et cognition – Catégories, prototypes, typicalité.* Paris, CNRS, p. 31-54.

Dubois, D., 2006, « Des catégories d'odorants à la sémantique des odeurs : une approche cognitive de l'olfaction ». *Terrain,* 47, p. 89-106.

Dubois, D., 2009, « Le sentir et le dire : définir l'objet et construire la démarche ». In Dubois, D. (éd.), *Le sentir et le dire.* Paris, L'Harmattan, p. 13-45.

Dubois, D., Bourgine, R., & Resche-Rigon, P., 1992, « Connaissance et expertises finalisées de divers acteurs économiques dans la catégorisation d'un objet perceptif ». *Intellectica,* 15, numéro spécial « Connaissances et rationalités : Questions ergonomiques et recherches cognitives », p. 241-271.

Dubois, D., Guastavino, C., & Raimbault, M., 2005, « Les catégories cognitives du bruit urbain : des discours aux indicateurs physiques ». *Acoustique et technique,* 39, p. 49-57.

Dubois, D., & Castellengo, M., 2006, *Du signal acoustique aux significations. Psychoacoustique et recherches cognitives*. Conférence présentée au 8e CFA (Congrès français d'acoustique), Tours.

Dubois, D., & Grinevald, C., 2003, « En voir de toutes les couleurs : processus de dénomination des couleurs et constructions cognitives ». In Vandeloise, C. (éd.), *Langues et cognition*. Paris, Hermès, p. 79-114.

Duda, R. O., 1996, « Auditory localization demonstrations ». *Acustica,* 82 (1), p. 346-355.

Dumoulin, N., 1996, *Étude de la qualité sonore d'un ensemble de clavecins.* DEA. Atiam (LAM), université Paris 6.

Dupin, F., 1971, « Lexique de la percussion ». *La Revue musicale,* numéro spécial, 284, p. 1-71. Paris, Richard Masse Éditeur.

Edelman, G. M., 1992, *Biologie de la conscience* (Gerschenfeld, A., trad.). Paris, Odile Jacob, 368 p.

Ehrenfels von, C., 1890, « Über Gestaltqualitäten ». *Vierteljahresschrift für Philosophie,* 14, p. 249-292.

Emery, E., 1998, *Temps et musique*, vol. I. *Temps et dialectique de la durée* ; vol. II. *Dialectique de la durée dans l'art musical*. Lausanne, L'Âge d'Homme, 696 p.

Engel, M., 1992, « La trompette marine – Réflexions d'un musicien, plus exactement d'un violoncelliste sur cet instrument ». In Collectif, *Amour et sympathie. Actes du colloque sur les instruments à cordes sympathiques*, Limoges, 28-29 novembre 1992. Ensemble baroque de Limoges, p. 63-97.

Engramelle, J., 1775, *La tonotechnie ou l'art de noter les cylindres*. Paris, chez P. M. Delaguette, libraire-imprimeur, 4 planches, 336 p.

Erickson, R., 1975, *Sound structure in music*. Berkeley, Univ. of California Press, 205 p.

Essid, S., 2005, *Classification automatique des signaux audio-fréquences : reconnaissance des instruments de musique.* Thèse de l'université Paris 6 [accessible sur http://perso.telecom-paristech.fr/~essid/papers/SE_PhD-05.pdf]

Fabre, B., & Castellengo, M., 1989, *Représentation de l'évolution du timbre des instruments de musique en fonction de la tessiture : application à l'orgue*. Conférence présentée à l'ICA (13th International Congress on Acoustics), Belgrade.

Fant, G., 1970, *Acoustic Theory of Speech Production*. La Haye/Paris (2e éd.), Mouton, 328 p.

Faure, A., & Mc Adams, S., 1997, *Comparaison de profils sémantiques et de l'espace perceptif de timbres musicaux.* Conférence présentée au 4e CFA (Congrès français d'acoustique), Marseille.

Faure, A., 2000, *Des sons aux mots : comment parle-t-on du timbre musical ?* Thèse d'université, EHESS, Paris.

Fechner, G. T., 1860, *Elemente der Psychophysik* (2 vol.). Leipzig, Breitkopf & Härtel, 336 p.

Fernando, N., 2007, « Échelles et modes : vers une typologie des systèmes scalaires ». In Nattiez, J. J. (éd.), *Musiques – Une encyclopédie pour le XXIe siècle*, vol. 5 *L'unité de la musique*. Paris, Actes Sud/Cité de la Musique, p. 945-979.

Féron, F.-X., 2006, *Des illusions auditives aux singularités du son et de la perception. L'impact de la psychoacoustique et des nouvelles technologies sur la création musicale au XXe siècle.* Thèse de l'université Paris 4.

Fichet, L., 1996, *Les théories scientifiques de la musique aux XIXe et XXe siècles*. Paris, Librairie philosophique J. Vrin, 381 p.

Fitch, W. T., & Fritz, J. B., 2006, « Rhesus macaques spontaneously perceive formants in conspecific vocalizations ». *J. Acoust. Soc. Am.,* 120 (4), p. 2132-2141.

Fletcher, H., & Sanders, L. C., 1967, « Quality of violin vibrato tones ». *J. Acoust. Soc. Am.,* 41 (6), p. 1534-1544.

Fletcher, H., 1934, « Loudness, pitch and the timbre of musical tones and their relation to the intensity, the frequency and the overtone structure ». *J. Acoust. Soc. Am.,* 6 (2), p. 59-69.

Fletcher, H., 1953 (2e éd.), *Speech and Hearing in Communication*. New York, D. Van Nostrand Company Inc., 461 p.

Fletcher, N. H., & Rossing, T. D., 1991, *The Physics of Musical Instruments*. New York, Springer-Verlag, 620 p.

Fontana, E., 2015, « Portraits of Johann George Tromlitz (1725-1805) ». *Musique-Images-Instruments,* 15. CNRS

Foch, A., 1934, *Acoustique*. Paris, Librairie Armand Colin, 208 p.

Fourier, J., 1822, *Théorie analytique de la chaleur*. Paris, Didot, 670 p.

Francès, R., 1984 (2e éd.), *La perception de la musique*. Paris, Librairie philosophique J. Vrin, 422 p.

François, J.-C., 1991, *Percussion et musique contemporaine*. Paris, Klincksieck, 300 p.

Fromentel, A., 2010, « La technologie audionumérique ». In Mercier, D. (éd.), *Le livre des techniques du son*, vol. 1. Paris, Dunod, p. 375-434.

Fürniss, S., 2000, « Cadres de quartes, fluctuations d'intervalles et mutations pentatoniques dans la polyphonie vocale aka ». *Musicae Scientiae,* numéro spécial « ESCOM-Discussion Forum 1 », p. 65-72.

Gaillard, P., 2000, *Étude de la perception des transitoires d'attaque des sons de steel-drums : particularités acoustiques, transformation par synthèse et catégorisation*. Thèse de l'université Toulouse 2 – Le Mirail.

Gaillard, P., 2009, « Laissez-nous trier ! » In Dubois, D., *Le sentir et le dire*. Paris, L'Harmattan, p. 189-210.

Gaillard, P., Castellengo, M., & Dubois, D., 2007, « L'apport de la catégorisation à l'étude du transitoire d'attaque du steel-drum ; contribution à la définition du timbre causal ». *Les Cahiers de la SQRM (Société québécoise de recherche en musique)*, 9 (1-2), p. 39-48.

Gaillard, P., Castellengo, M., & Legros, C., 2000, *Modification de certaines caractéristiques physiques des sons de steel-drums en vue de la réalisation de tests de psychoacoustiques*. Conférence présentée au 5e CFA (Congrès français d'acoustique), Lausanne.

Ganassi, S., 1535, *Opera Intitulata Fontegara*. Bologne, Forni Editore (fac-similé, 1969), 160 p.

Garcia (Jr), M., 1884 (8e éd.), *École de Garcia, Traité complet de l'art du chant*. Paris, Heugel, 100 p.

Garnault (Dr), P., 1895, *Cours théorique et pratique de physiologie, Hygiène et thérapeutique de la voix parlée et chantée, Hygiène et maladies du chanteur et de l'orateur*. Paris, Maloine, 464 p.

Garnier, M., 2003, *Approche de la qualité vocale dans le chant lyrique : perception, verbalisation et corrélats acoustiques*. DEA. Atiam (LAM), université Paris 6.

Garnier, M., Dubois, D., Henrich, N., Castellengo, M*., et al.*, 2005, « Étude de la qualité vocale dans le chant lyrique ». *Scolia,* 20, p. 151-169.

Garnier, M., Henrich, N., Castellengo, M., Sotiropoulos, D*., et al.,* 2007, « Characterisation of Voice Quality in Western Lyrical Singing: from Teacher's Judgements to Acoustic Descriptions ». *Journal of Interdisciplinay Music Studies (JIMS),* 1 (2), p. 62-91.

Gellé (Dr), M. E., 1899, *L'audition et ses organes*. Paris, Félix Alcan, 326 p.

Gérard-Ardenois, A., 2002, « Primera aproximación a la acùstica de la "tarka" ». *Revista Boliviana de Fisica* (8), p. 42-51.

Gherghinoiu, C., Besnainou, C., & Castellengo, M., 2005, *Comparative study made on timbre perception.* Conférence présentée à l'EAA Forum Acusticum Conference, Budapest.

Gibson, J. J., 1979, *The Ecological Approach to Visual Perception*. Boston, Houghton Mifflin, 325 p.

Gilbert, J., Kergomard, J., & Kientzy, D., 1987, « Le saxophone ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 115. LAM, université Paris 6.

Goad, P. J., 1994, *Timbral sharpness and modulations in frequency and amplitude: implications for the fusion of musical sounds*, Phd., University of Washington, 156 p.

Goad, P. J., & Castellengo, M., 1997. *Étude comparée de la qualité sonore du jeu de cromorne dans quatre orgues classiques : évaluation psychoacoustique de la nasalité.* Conférence présentée au 4e CFA (Congrès français d'acoustique), Marseille.

Goad, P. J., & Keefe, D. H., 1992, « Timbre discrimination of Musical instruments in a Concert Hall ». *Music Perception,* 10 (1), p. 43-62.

Godøy, R. I., 1993. *Formalization and epistemology.* Phd, faculté de philosophie historique, Oslo.

Godøy, R. I., 1997, « Knowledge in Music Theory by Shapes of Musical Objects and Sound-Producing Actions ». In Leman (éd.), *Music, Gestalt, and Computing*. Berlin, Springer, p. 89-102.

Godwin, J., 1980, *Athanasius Kircher – Un homme de la Renaissance à la quête du savoir perdu* (Matton, S., trad.). Londres, Thames & Hudson (J.-J. Pauvert), 96 p.

Gombrich, E. H., 2002, *L'art et l'illusion. Psychologie de la représentation picturale* (Durand, G., Haddad, N., & Préface, trad. 6e éd.). Washington/Paris, Phaidon Press Ltd, 386 p.

Gordon, J. W., & Grey, J. M., 1978, « Perception of spectral modifications on orchestral instrument tones ». *Computer Music Journal,* 2 (1), p. 24-31.

Goto, M., Haschiguchi, H., Nishimura, T., & Oka, R., 2003, *RWC Music Database: Music Genre Database and Musical Instrument Sound Database.* Conférence présentée à l'ISMIR, Baltimore.

Grey, J. M., & Gordon, J. W., 1978, « Perceptual effects of spectral modifications on musical timbres ». *J. Acoust. Soc. Am.,* 63 (5), p. 1493-1500.

Grey, J. M., & Moorer, J. A., 1977, « Perceptual evaluations of synthesized musical instrument tones ». *J. Acoust. Soc. Am.,* 62 (2), p. 454-462.

Grey, J. M., 1977, « Categorical versus continuous Perception of Musical Timbre ». In *Experiments in timbre perception*. États-Unis, Stanford University Press, p. 142-157.

Grey, J. M., 1977, « Multidimensional perceptual scaling of musical timbres ». *J. Acoust. Soc. Am.,* 61 (5), p. 1270-1277.

Grey, J. M., 1978, « Timbre discrimination in musical patterns ». *J. Acoust. Soc. Am.,* 64 (2), p. 467-472.

Grisey, G., 1991, « Structuration des timbres dans la musique instrumentale ». In Barrière, J.- B. (éd.), *Le timbre – Métaphore pour la composition.* Paris, Christian Bourgois/ IRCAM, p. 352-385.

Grisey, G., 2008, *Écrits* (édition établie par Guy Lelong). Paris, MF, 375 p.

Guastavino, C., 2009, « Validité écologique des dispositifs expérimentaux. Pour qui ? Pour quoi ? Pour quoi faire ? » In Dubois, D. (éd.), *Le sentir et le dire.* Paris, L'Harmattan, p. 233-252.

Guillaume, P., 1931, *Psychologie*. Paris, Félix Alcan, 388 p.

Guillaume, P., 1937, *La psychologie de la forme.* Paris, Champs-Flammarion, 234 p.

Guillaume, P., 1960, *Manuel de psychologie*. Paris, PUF, 315 p.

Guillemin, A., 1875, *Le Son. Notions d'acoustique physique et musicale*. Paris, Hachette et Cie, 268 p.

Guyot, F., 1996, *Étude de la perception sonore en termes de reconnaissance et d'appréciation qualitative : une approche par la catégorisation*. Thèse de l'université du Maine et UPMC, Le Mans.

Guyot, F., Castellengo, M., & Fabre, B., 1997, « Étude de la catégorisation d'un corpus de bruits ». In Dubois, D. (éd.), *Catégorisation, représentation et systèmes symbolique.* Paris, Kimé, p. 45-58.

Guyot, F., Castellengo, M., Vogel, C., & Maffiolo, V., 1996, « Une méthode d'étude de la qualité acoustique des sons réels complexes ». *Acoustique et Techniques,* 7, p. 23-26.

Habermann, G., 1978, *Stimme und Sprache. Eine Einführung in ihre Physiologie und Hygiene*. Stuttgart, Thieme, 27 p.

Hajda, J. M., 2007, « The effect of dynamic acoustical features on musical timbre ». In Collectif – Beauchamp, J. W. (éd.), *Analysis, Synthesis, and Perception of Musical Sounds – The sound of music*. New York, Springer, p. 250-271.

Hajda, J. M., Kendall, R. A., Carterette, E. C., & Harschberger, M. L., 1994, « Methodological issues in Timbre research ». In Deliège, I., & Sloboda, J. (éd.), *Perception and Cognition of Music.* New York, Psychology Press, p. 253-306.

Handel, S., & Erickson, M. L., 2001, « A Rule of Thumb: The Bandwidth for Timbre Invariance is One Octave ». *Music Perception,* 19 (1), p. 121-126.

Handel, S., 1989, *Listening – An Introduction to the perception of auditory events*. Cambridge, MIT Press, 597 p.

Handel, S., 2006, *Perceptual Coherence: Hearing and Seeing*. New York, Oxford Univ. Press, 469 p.

Harvey, J., 1991, « Le miroir de l'ambiguïté ». In Barrière, J.-B. (éd.), *Le timbre – Métaphore pour la composition.* Paris, Christian Bourgois/IRCAM, p. 454-466.

Haynes, B., 2002, *A history of performing pitch. The story of "A"*. Lanham (États-Unis), The Scarecrow Press Inc., 569 p.

Helmholtz (von), H., 1863, *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Réimpression à l'identique (1981) de l'édition Vieweg und Sohn, Brunswick, 1863. Francfort-sur-le-Main, Minerva Verlag GMBH, 600 p.

Helmholtz (von), H., 1874, *Théorie physiologique de la musique fondée sur l'étude des sensations auditives* (Guéroult, F., trad.). Paris, Masson (fac-similé J.-J. Gabay, 1990), 544 p.

Henrich, N., Bezard, P., Expert, R., Garnier, M., *et al.*, 2007, *Perception and verbalisation of voice quality in Western lyrical singing: contribution of a muldisciplinary research group*. Conférence présentée au CIM07 (Conference on Interdisciplinary Musicology), Tallin (Estonie).

Henrich, N., Roubeau, B., & Castellengo, M., 2003, *On the use of electroglottography for characterisation of the laryngeal mechanisms.* Conférence présentée au SMAC03 (Stockholm Music Acoustics Conference), Stockholm.

Henrich, N., Smith, J., & Wolfe, J., 2011, « Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones ». *J. Acoust. Soc. Am.,* 129, p. 1024-1035.

Henrich-Bernardoni, N., 2014, « Vibrations et résonances en voix chantée ». In Henrich-Bernardoni, N. (éd.), *La voix chantée, entre sciences et pratiques*. Paris/Bruxelles, De Boeck/Solal, p. 3-20.

Henry, P., 1979, *Journal de mes sons*. Paris, Séguier, 23 p.

Hirano, M., 1982, « The role of the layer structure of the vocal fold in register control ». In *Vox Humana – Studies presented to Aato Sonninen*. Jyväskylä (Finlande), Institute of Finnish Language and Communication, University of Jyväskylä, p. 50-62.

Hirano, M., Hibi, S., & Hagino, S., 1995, « Physiological Aspects of Vibrato ». In Dejonckere, P. H., Hirano, M., & Sundberg, J. (éd.), *Vibrato*. San Diego, Singular Publishing Group, p. 9-33.

Holland, J., 1978, *Percussion* (1re éd.). Londres, MacDonald and Jane's Publishers Limited, 283 p.

Hotteterre-le Romain, J. M., 1707, *Principes de la flûte traversière, ou flûte d'Allemagne, de la flûte à bec, ou flûte douce et du haut-bois*. Amsterdam, Estienne Roger, 47 p.

Hourdin, C., 1995, *Étude psychophysique du timbre – Application au codage et à la synthèse des sons en musique*. Thèse de l'université Paris 11, Orsay.

Houtsma, A., 1995, *Tuning of orchestral chimes: pitch and consonance.* Conférence présentée à l'ISMA (International Symposium on Musical Acoustics), Dourdan (Essonne).

Hugonnet, C., & Walder, P., 1995, *Théorie et pratique de la prise de son stéréophonique*. Paris, Eyrolles, 251 p.

Husserl, E., 1964 (6e éd.), *Leçons pour une phénoménologie de la conscience intime du temps* [cours année 1904-1905] (Dussort, H., trad.). Paris, PUF, 205 p.

Husson, R., 1961, *Physiologie de la phonation*. Paris, Masson, 590 p.

Huyghe, R., 1971, *Formes et forces, de l'atome à Rembrandt*. Paris, Flammarion, 444 p.

Iverson, P., & Krumhansl, C., 1993, « Isolating the dynamic attributes of musical timbre ». *J. Acoust. Soc. Am.,* 94 (5), p. 2595-2603.

Jakobson, R., 1963, *Essais de linguistique générale* (Ruwet, N., trad.). Paris, Les Éditions de Minuit, 260 p.

Janson, E. V., 1990, *Violin timbre and the picket fence*, STL-QPSR 2-3/1990, Stockholm, KTH.

Kanizsa, G., 1998, *La grammaire du voir. Essais sur la perception*. Paris, Diderot, 400 p.

Kastner, G., 1837, *Traité général d'instrumentation*. Paris, P. Philip.

Katz, D., 1955, *Introduction à la psychologie de la forme* (David, M., & Voute, S., trad.). Paris, Librairie Marcel Rivière et Cie, 188 p.

Kendall, R. A., 2002, *Musical timbre beyond a single note, II : interactions of pitch chroma and spectral centroid*. Conférence présentée à la 7th ICMPC (International Conference on Music Perception and Cognition), Sydney.

Kergomard, J., 1976, « Le basson ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 82-83. LAM, université Paris 6.

Koechlin, C., 1935, *Traité de l'orchestration* (vol. 1). Paris, Éditions Max Eschig (fac-similé, 1954), 322 p.

Koenig, R., 1882, *Quelques expériences d'acoustique*. Paris, 27 quai d'Anjou, 248 p.

Koenig, R., 1889, *Catalogue des appareils d'acoustique*. Chez l'auteur, Paris, 27 quai d'Anjou, 100 p.

Koenig, W., Dunn, H. K., & Lacy, L. Y., 1946, « The sound Spectrograph ». *J. Acoust. Soc. Am.,* 18 (1), p. 19-49.

Koffka, K., 1935, *Principles of Gestaltpsychology*. New York, Harcourt, 720 p.

Köhler, W., 1964, *Psychologie de la forme – Introduction à de nouveaux concepts en psychologie* (Bricianer, S., trad.). Paris, Gallimard, 373 p.

Kondratov, A., 1966, *Sons et signes* (Sokolov, I., trad.). Moscou, Mir, 284 p.

Krimphoff, J., Mc Adams, S., & Windsberg, S., 1994, *Caractérisation du timbre des sons complexes : Analyses acoustiques et quantification psychophysique.* Conférence présentée au 3e CFA (Congrès français d'acoustique), Toulouse.

Krumhansl, C., 1989, « Why is musical timbre so hard to understand? » In Nielzen, S., & Olson, O. (éd.), *Structure and perception of electroacoustic sound and music*. Amsterdam, Elsevier, p. 43-53.

Labussière, A., 2007, « Geste et structure modale dans le chant traditionnel à voix nue ». In Nattiez, J. J. (éd.), *Musiques – Une encyclopédie pour le XXIe siècle*, vol. 5 *L'unité de la musique*. Paris, Actes Sud/Cité de la Musique, p. 980-1024.

Lähdeoja, O., Navarret, B., Quintans, S., & Sèdes, A., 2013, « La guitare électrique comme instrument augmenté et outil de création musicale ». In Castellengo, M. & Genevois, H. (éd.), *La musique et ses instruments*. Paris, Delatour, p. 313-329.

Lakatos, S., 2000, « A common perceptual space for harmonic and percussive timbres ». *Perception & Psychophysics,* 62 (7), p. 1426-1439.

Lakoff, G., 1987, *Women, Fire, and Dangerous Things – What Categories Reveal about the Mind*. Chicago, The University of Chicago Press, 614 p.

Lattard, J., 1988, *Gammes et tempéraments musicaux*. Paris, Masson, 130 p.

Lattard, J., 1997, *Musique : gammes et tempéraments – De Pythagore aux simulations informatiques*. Paris, Diderot Multimedia, 219 p.

Le Huche, F., & Allali, A., 1984, *La voix* (vol. 1). Paris, Masson, 223 p.

Le Nindre, B., 2006, *Typage sonore identitaire des bruits automobiles : application au cas du véhicule sportif*. Thèse d'université, ENSAM, Paris.

Lecocq, P., 1993, « La structure de la mémoire humaine ». In Weil-Barais, A. (éd.), *L'homme cognitif*. Paris, PUF, p. 351-369.

Legent, F., Perlemuter, L., & Vandenbrouck, C., 1968, *Cahiers d'anatomie ORL.* Paris, Masson et Cie, 135 p.

Legros, H., 1972, « Le tempérament ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 61. LAM, université Paris 6.

Lehr, A., 1976, *Leerboek der Campanologie – Een muzikaal-technische verhandeling over luidklokken en beiaarden*. Asten (Pays-Bas), Nationaal Beiaardmuseum, 210 p.

Leipp, E., 1965, in « La vielle, le monocorde vietnamien, la cithare à 16 cordes ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 12. LAM, université Paris 6.

Leipp, E., 1965, « Le problème de la perception des signaux acoustiques par effet de contraste : les signaux d'avertissement, les sifflets ». *Annales des télécommunications,* 20 (5-6), p. 103-118.

Leipp, E., 1965, « Les champs de liberté des instruments de musique ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 10. LAM, université Paris 6.

Leipp, E., 1967, « Mécanique et acoustique de l'appareil phonatoire ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 32. LAM, université Paris 6.

Leipp, E., 1968, « Structure physique et contenu sémantique de la parole ». *Revue d'Acoustique,* 3-4, p. 259.

Leipp, E., 1974, « Les sons nouveaux avec instruments de musique mécaniques ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 74. LAM, université Paris 6.

Leipp, E., 1976, *Acoustique et musique* (2e éd.). Paris, Masson, 337 p.

Leipp, E., 1977, « L'intégrateur de densité spectrale IDS et ses applications ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 94. LAM, université Paris 6.

Leipp, E., 1977, *La machine à écouter – Essai de psycho-acoustique*. Paris, Masson, 260 p.

Leipp, E., 2011, *Acoustique et musique* (fac-similé de la 4e éd., 1984). Paris, Presses des Mines, 374 p.

Leipp, E., & Castellengo, M., 1964, « Qu'est-ce qu'un son de clarinette ? – Was ist ein Klarinettenklang? » *Das Musikinstrument*, 1964 (8).

Leipp, E., & Castellengo, M., 1977, « Du diapason et de sa relativité ». *La Revue musicale,* n° 294, (39 p.). Paris, Richard Masse Éditeur.

Leipp, E., & Trân Van Khe, 1967, « Réflexions sur les problèmes du rythme dans les musiques orientales ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 29. LAM, université Paris 6.

Leipp, E., Castellengo, M., & Agostini, F., 1969, « Le nouveau diapason électronique à l'Opéra ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 40. LAM, université Paris 6.

Leipp, E., Maurin, M., & Genet-Varcin, E., 1976, « La cloche-diapason du Conservatoire national de musique de Paris ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 88. LAM, université Paris 6.

Leipp, E., Sapaly, J., & Castellengo, M., 1971, « Présentation de notre orgue expérimental, le Cantor ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 56. LAM, université Paris 6.

Leipp, W., Castellengo, M., Liénard, J.-S., Sapaly, J.*, et al.,* 1971, « Les recherches sur la parole au laboratoire d'Acoustique ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 53. LAM, université Paris 6.

Lévêque, Y., 2012, *Le lien perception-production en voix chantée : place des représentations motrices*. Thèse de l'université d'Aix-Marseille, Marseille.

Lévêque, Y., 2013, « La voix chantée : explorer la boucle audio-vocale en musique ». *La lettre des neurosciences,* n° 45, p. 13-15.

Levin, T. C., & Edgerton, M. E., 1999, « The throat singers of Tuva ». *Scientific American,* septembre, p. 80-87.

Lévi-Strauss, C., 1964, *Le cru et le cuit*. Paris, Plon, 402 p.

Levitin, D. J., & Rogers, S. E., 2005, « Absolute pitch : perception, coding, and controversies ». *Trends in Cognitive Sciences,* 9 (1), p. 26-33.

Lévy, F., 2008, *Mutations (1969) de Jean-Claude Risset : une écriture déjouée par des paradoxes*, Paris, INA/CDMC. http://www.fabienlevy.net/Documents/pdf/ArticleRisset-Polychrone.pdf

Lichte, W. H., 1941, « Attributes of complex tones ». *Journal of experimental Psychology,* 28 (6), p. 455-480.

Liénard, J.-S., 1977, *Les processus de la communication parlée – Introduction à l'analyse et la synthèse de la parole*. Paris, Masson, 189 p.

Liénard, P., 2010, « Acoustique fondamentale ». In Mercier, D. (éd.), *Le livre des techniques du son*, vol. 1 (4e éd.). Paris, Dunod, p. 1-43.

Ligeti, G., 2001, « Musique et Technique. Expériences personnelles et considérations subjectives » (traduction d'un texte écrit en 1980). In Ligeti, G. (éd.), *Neuf essais sur la musique.* Genève, Contrechamps, p. 181-206.

Lindley, M., 1984, *Lutes, viols & temperaments*. Cambridge, Cambridge University Press, 134 p.

Lindsay, P., & Norman, D., 1980, *Traitement de l'information et comportement humain – Une introduction à la psychologie*. Montréal, Vigot/Études Vivantes, 754 p.

Locher, J. L., Broos, C. H. A., Escher, M. C., Locher, G. W.*, et al.*, 1972, *Le monde de M. C. Escher* (Renaud, J., trad.). Paris, Éditions du Chêne, 270 p.

Lortat-Jacob, B., 1998, *Chants de Passion. Au cœur d'une confrérie de Sardaigne* (livre avec CD). Paris, Éditions du Cerf, 343 p.

Loulié, E., 1696, *Éléments ou principes de musique, mis dans un nouvel ordre*. Paris, Christophe Ballard, 96 p.

Luce, D., & Clark, M. J., 1965, « Durations of Attack Transients of Nonpercussive Orchestral Instruments ». *J. Audio Eng. Soc.,* 13, p. 194-199.

Luce, D., & Clark, M. J., 1967, « Physical correlates of Brass-Instrument tones ». *J. Acoust. Soc. Am.,* 42 (6), p. 1232-1243.

Mach, E., 1885, « Zur Analyse der Tonempfindungen ». *Sitzungbericht Kaiserlich Akademie Wissenschaft,* 92 (2), p. 1283-1289.

Madelaine (de la), S., vers 1851, *Théories complètes du chant*. Paris, Amyot, rue de la Paix, 412 p.

Maffiolo, V., 1999, *De la caractérisation sémantique et acoustique de la qualité sonore de l'environnement urbain*. Thèse de l'université du Maine, Le Mans.

Maidhof, C., & Koelsch, S., 2011, « Effects of selective attention on syntax processing in music and language ». *Journal of Cognitive Neuroscience,* 23 (9), p. 2252-2267.

Maisonneuve, S., 2009, *L'invention du disque (1877-1949).* Paris, Éditions des Archives Contemporaines, 280 p.

Manoury, P., 1991, « Les limites de la notion “de timbre” ». In Barrière, J.-B. (éd.), *Le timbre – Métaphore pour la composition.* Paris, Christian Bourgois/IRCAM, p. 293-300.

Manoury, P., 1998, *La note et le son, écrits et entretiens (1981-1998).* Paris, L’itinéraire/L’Harmattan, 421 p.

Marage (Dr), R., 1911, *Petit manuel de physiologie de la voix à l’usage des chanteurs et des orateurs.* Paris, chez l’auteur (accessible sur openlibrary.org/books/OL24636918M), 204 p.

Marandas, E., Mostis, K., & Gibiat, V., 1998, *Physique et perception de l’accord des instruments à clavier. Du clavecin au piano* (CD d’exemples sonores). Conférence présentée au colloque « Acoustique et instruments anciens. Factures, Musique et Sciences », Paris, Cité de la Musique.

Martin, K. D., 1998, *Toward automatic sound source recognition : identifying musical instruments*. Conférence présentée au NATO Advanced Study Institute on Computational Hearing, [accessible sur http://alumni.media.mit.edu/~kdm/], Il Ciocco (Italie).

Martin, K. D., 1999, *Sound-Source Recognition : A theory and Computational Model.* Phd, M.I.T., Cambridge, M.A. [http://96db.com/static/kdm-phdthesis.pdf].

Mathews, M., 1977, « Analysis and synthesis of timbres ». In *Music, Room, Acoustics.* Stockholm, Royal Swedish Academy of Music, vol. 17, p. 4-18.

McAdams, S., 1994, « La reconnaissance des sources et d’événements sonores » (Henry, J., trad.). In McAdams, S. & Bigand, E. (éd.), *Penser les sons : psychologie cognitive de l’audition*. Paris, PUF, p. 155-213.

McAdams, S., & Cunibile, J. C., 1992, « Perception of timbral analogies ». *Philosophical Transactions of the Royal Society of London,* 336 (London Series B 1992), 11 pages.

McAdams, S., Depalle, P., & Clarke, E., 2004, « Analyzing musical sound ». In Clarke, E. & Cook, N. (éd.), *Empirical Musicology. Aims, Methods, and Prospects*. Oxford, Oxford University Press, p. 157-196.

Merleau-Ponty, M., 1962, « Un inédit de Maurice Merleau-Ponty publié par Martial Guéroult ». *Revue de métaphysique et de morale,* 4, p. 401-409.

Mersenne, M., 1636, *Harmonie universelle contenant la théorie et la pratique de la musique* Paris, CNRS (fac-similé, 1963), vol. 1, 624 p., vol. 3, 640 p.

Metfessel, M., 1928, *Phonophotography in Folk Music – American negro songs in new notation.* Chapel Hill, The University of North Carolina Press, 181 p.

Meyer, C., 1996, *Mensura monocordi : la division du monocorde – Textes des IXe-XVe siècles*. Paris, Société française de musicologie/Klincksieck, 462 p.

Meyer, J., 1993, « The Sound of the Orchestra ». *J. Audio Eng. Soc.,* 41 (4), p. 203-213.

Miller, D. C., 1916, *The science of musical sounds*. New York, The Macmillan Company, 286 p.

Miller, J. R., & Carterette, E. C., 1975, « Perceptual space for musical structures ». *J. Acoust. Soc. Am.,* 58 (3), p. 711-720.

Miller, R., 1990, *La structure du chant – Pédagogie systématique de l’art du chant* (Gouëlou, J.-M., trad.). Paris, Éditions IPMC, 395 p.

Miyazaki, Ken’ichi, 1988, « Musical pitch identification by absolute pitch possessors ». *Perception & Psychophysics,* 44 (6), p. 501-512.

Miyazaki, Ken’ichi, 2004, « How well do we understand absolute pitch ». *Acoust. Sci. & Tech.,* 25 (6), p. 426-432.

Moles, A., 1952, « Étude et représentation de la note complexe en acoustique musicale ». *Cahiers d’acoustique du GALF,* 46, p. 430-438.

Moles, A., 1960, *Les musiques expérimentales. Revue d’une tendance importante de la musique contemporaine*. Paris, Éditions du Cercle d’art contemporain, 161 p.

Moles, A., 1972, *Théorie de l’information et perception esthétique*. Paris, Denoël/Gonthier, 327 p.

Moles, A., 1990, *Les sciences de l’imprécis*. Paris, Seuil, 253 p.

Molino, J., 1999, « La musique et l’objet ». In *Ouïr, entendre, écouter, comprendre après Schaeffer*. Paris, Buchet-Chastel/INA, p. 119-136.

Montchalin (de), G., 1995. *Étude de la qualité sonore d’un ensemble de guitares par une méthode de catégorisation.* Mémoire de stage, École polytechnique, Université Paris 6, Paris.

Monteil, M., 2010, *Vibrations non linéaires de steel-drums. Caractérisation expérimentale et modèle phénoménologique*. DEA. Atiam (CNAM), université Paris 6.

Moore, B. C. J., 2013 (6e éd.), *An introduction to the Psychology of Hearing*. Brill, 420 p.

Mori, T., 2000, *Ein Vergleich der qualitätsbestimmenden Faktoren von Klavier und Flügel.* Thèse de doctorat, Tech. Univ. Carolo-Wilhelmina, Brunswick (Allemagne).

Morse, P. M., 1948 (2e éd.), *Vibration and Sound*. New York, McGraw-Hill Book Company Inc., 468 p.

Morse, P. M., & Ingard, K. U., 1968, *Theoretical Acoustics*. New York, McGraw-Hill Book Company Inc., 927 p.

Mukherjee, P., 2010, *Les échelles fondamentales de la musique indienne du Nord et du Sud*. Paris, Éditions Publibook Université, 280 p.

Musson-Gonneaud, V., & Besnainou, C., 2013, « Les harpions, questions organologiques et musicales : quel réglage pour quel usage sur les harpes anciennes aujourd'hui ? » In Castellengo, M., & Genevois, H. (éd.), *La musique et ses instruments* (+ DVD). Paris, Delatour, p. 37-66.

Nattiez, J.-J., 2007, « Le timbre est-il un paramètre secondaire ? » *Les cahiers de la SQRM* (Société québécoise de recherche en musique), 9 (1-2), p. 13-24.

Neisser, U., 1967, *Cognitive psychology*. New York, Appleton-Century-Crofts, 351 p.

Neisser, U., 1987, *Concepts and conceptual development.* Cambridge, Cambridge University Press, 317 p.

Nicolas, S., 2002, « La fondation de la psychophysique de Fechner : des présupposés métaphysiques aux écrits scientifiques de Weber ». *L'année psychologique,* n° 102 (n° 2), p. 255-298 [Site Persée http://www.persee.fr/web/revues].

Nicolas, S. (éd.), 2003, *La psychologie cognitive*. Paris, Armand Colin.

Ninio, J., 1989, *L'empreinte des sens – La raison perceptive*. Paris, Odile Jacob, 304 p.

Noistermig, M., Carpentier, T., & Warusfel, O., 2012, « Dispositif de spatialisation sonore 3D à l'espace de projection de l'IRCAM – Un réseau de 345 haut-parleurs pour une restitution par WFS et HOA ». *Acoustique et Techniques,* 71, p. 30-39.

Olson, H. F., 1967 (2e éd.), *Music, Physics and Engineering*. New York, Dover Publications Inc., 460 p.

Ormezzano, Y., 2000, *Le guide de la voix*. Paris, Odile Jacob, 432 p.

Pascal, B., 1897, *Pensées*. http://www.penseesdepascal.fr/Misere/Misere.php. CNRS, Bnf, Clermont université.

Patil, K., Pressnitzer, D., Shamma, S., & Elhilali, M., 2012, « Music in our ears: the biological bases of musical timbre perception ». *PLoS Computational Biology,* 8 (11), p. e1002759.

Patterson, B., 1974, « Musical dynamics ». *Scientific American,* 231 (5), p. 78-95.

Patterson, R. D., Smith, D. R. R., Dinther, R. v., & Walters, T. C., 2010, « Size information in the production and perception of communication sounds ». In Yost, W. A., Popper, A. N. & Fay, R. R. (éd.), *Auditory perception of sound sources*. New York, Springer, LLC, p. 43-75.

Peeters, G., 2003, *Automatic classification of large musical instrument databases using hierarchical classifiers with inertia ratio maximization*. Conférence présentée à la 115th AES Convention, New York.

Peeters, G., Giordano, B., Susini, P., Misdariis, N.*, et al.,* 2011, « The timbre toolbox: extracting audio descriptors from musical signals ». *J. Acoust. Soc. Am.,* 130 (5), p. 2902-2916.

Penrose, L., & Penrose, R., 1958, « Impossible objects; A special type of visualisation », *British Journal of Psychology*. 49 (1), p. 31-33.

Perry, D. W., Zatorre, R. J., Petrides, M., Alivisatos, B.*, et al.,* 1999, « Localization of cerebral activity during simple singing ». *Neuroreport,* 10 (18), p. 3979-3984.

Peterson, G. E., & Barney, H. L., 1952, « Control methods used in a study of the vowels ». *J. Acoust. Soc. Am.,* 24 (2), p. 175-184.

Petitot, J., 1989, « Forme », in *Encyclopædia Universalis*. Chicago, Encyclopædia Britannica Inc., p. 712-728.

Pickles, J. O., 1982, *An Introduction to the Physiology of Hearing*. Orlando, Academic Press Inc., 341 p.

Pierce, J. R., 1984, *Le son musical – Musique, acoustique et informatique* (+ disque) (Berquier, F., trad.). Paris, Pour la Science, diffusion Belin, 242 p.

Piéron, H., 1945, *La sensation guide de vie*. Paris, NRF, Gallimard, 420 p.

Piéron, H., Chocholle, R., & Leplat, J., 1969 (2e éd.), *Traité de psychologie expérimentale – Sensation et motricité*, vol. 2. Paris, PUF, 171 p.

Plomp, R., 1976, *Aspects of tone sensation – A psychophysical study*. New York, Academic Press, 167 p.

Plomp, R., 2002, *The intelligent ear – On the nature of sound perception*. London, Laurence Erlbaum Associates Publ., 174 p.

Poizat, M., 1991, *La voix du diable – La jouissance lyrique sacrée*. Paris, Métailié, 249 p.

Pollard, H. F., 1988, « Feature analysis of musical sounds ». *Acustica,* 65, p. 232-244.

Pollard, H. F., 1999, « Tonal portrait of a pipe organ ». *J. Acoust. Soc. Am.,* 106 (1), p. 360-370.

Potard, J.-Y., 1987, *Interaction Matériau-Organisation*, rapport de recherche, Paris, IRCAM.

Praetorius, M., 1619, *Syntagma musicum,* vol. II *De Organographia*. Wolfenbüttel (réédition, Bärenreiter, 1958), 236 p. et 36 planches.

Prame, E., 1997, « Vibrato extent and intonation in professional western lyric singing ». *J. Acoust. Soc. Am.,* 102 (1), p. 616-621.

Pressnitzer, D., & Patterson, R., 2001, « Distortion products and the perceived pitch of harmonic complex tones ». In Breebart, D., Houtsma, A. J. M., Kohlrausch, A., Prijs, V. F., & Schoonoven, R. (éd.), *Physiological and Psychophysical Bases of Auditory Function*. Maastricht (Pays-Bas), Shaker Publishing BV, p. 97-104.

Pressnitzer, D., 1998, *Perception de rugosité psychoacoustique : d'un attribut élémentaire de l'audition à l'écoute musicale*. Thèse de l'université Paris 6.

Pressnitzer, D., Patterson, R., & Krumbholz, K., 2001, « The lower limit of melodic pitch ». *J. Acoust. Soc. Am.,* 109 (5), p. 2074-2084.

Proust, M., 1994, *Le Côté de Guermantes*. Paris, Classiques Garnier.

Psychoyou, T., 2003, *L'évolution de la pensée théorique, en France, de Marin Mersenne à Jean-Philippe Rameau*. Vol. 1 *Les modèles de la théorie musicale*. Vol. 2 *Écrits concernant la musique en France : 1623-1722*. Thèse de l'université de Tours, 389 p. et 757 p.

Radau, R., 1880, *Acoustique*. Paris, Librairie Hachette, 309 p.

Rasch, R., & Plomp, R., 1982, « The perception of musical tones ». In Deutsch, D. (éd.), *The psychology of Music* (2e éd., 1999). San Diego, Academic Press, p. 89-112.

Rameau, J.-P., 1722, *Traité de l'harmonie réduite à ses principes naturels*. Paris, J. B. C. Ballard, 432 p.

Risset, J.-C., 1966, *Computer study of trumpet tones*, Murray Hill (New Jersey), Bell Telephone Laboratories, 72 p.

Risset, J.-C., 1967, *Sur l'analyse, la synthèse et la perception des sons, étudiées à l'aide de calculateurs électroniques*. Thèse d'université, faculté des Sciences d'Orsay.

Risset, J.-C., 1969, *An introductory catalogue of computer-synthesized sounds*. Murray Hill (New Jersey), Bell Telephone Laboratories.

Risset, J.-C., 1971, *Paradoxe de hauteur : le concept de hauteur sonore n'est pas le même pour tout le monde.* Conférence présentée à l'ICA (7th International Congress on Acoustics), Budapest.

Risset, J.-C., 1978, « Musical acoustics ». In Carterette, E. C, et Friedman, M. P. (éd.), *Handbook of Perception*, vol. 4 *Hearing*. New York, Academic Press, p. 521-564.

Risset, J.-C., 1978, *Paradoxes de hauteur*, rapport de recherche n° 10, Paris, IRCAM.

Risset, J.-C., 1991, « Timbre et synthèse des sons ». In Barrière, J.- B. (éd.), *Le timbre, métaphore pour la composition*. Paris, IRCAM/Christian Bourgois, p. 239-271.

Risset, J.-C., 1994, « Quelques aspects du timbre dans la musique contemporaine ». In Zenatti, A. (éd.), *Psychologie de la musique*. Paris, PUF, p. 87-114.

Risset, J.-C., 2004, « Timbre ». In Nattiez, J.-J. (éd.), *Musiques, Une encyclopédie pour le XXe siècle*, vol. 2 *Les savoirs musicaux.* Paris, Actes Sud/Cité de la Musique, p. 134-161.

Risset, J.-C., 2014, *Composer le son. Repères d'une exploration du monde sonore numérique*. Paris, Hermann, 442 p.

Risset, J.-C., & Wessel, D., 1982, « Exploration of timbre by analysis and synthesis ». In Deutsch, D. (éd.), *Psychology of Music*. New York, Academic Press, p. 25-57.

Ritsma, R. J., 1962, « Existence region of tonal residue ». *J. Acoust. Soc. Am.,* 34 (9), p. 1224-1229.

Ritsma, R. J., 1967, « Frequencies dominant in the perception of the pitch of complex sounds ». *J. Acoust. Soc. Am.,* 42 (1), p. 191-198.

Rosch, E., & Lloyd, B. (éd.), 1978, *Cognition and categorization*. Hillsdale (New Jersey), L. Erlbaum.

Rosch, E., 1976, « Classifications d'objets du monde réel : origines et représentations dans la cognition ». *Bull. de Psychologie* (numéro spécial : « La mémoire sémantique »), p. 242-250.

Rossing, T. D., 1990 (2e éd.), *The Science of Sound*. Boston, Addison-Wesley Publishing Company, 686 p.

Rossing, T. D., Moore, R., & Wheeler, P., 2002 (3e éd.), *The Science of Sound*. Boston, Addison-Wesley Publishing Company, 783 p.

Roubeau, B., Henrich, N., & Castellengo, M., 2009, « Laryngeal vibratory mechanisms : the notion of vocal register revisited ». *Journal of Voice,* 23, p. 425-438.

Rousseau, J.-J., 1751, Article « Son, en Musique ». In Diderot, D. & d'Alembert, J. L. R. (éd.), *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers*. Paris, Le Breton, t. 15, p. 345.

Rousseau, J.-J., 1768, *Dictionnaire de musique*. Paris, La Veuve Duchesne, 548 p.

Roy, S., 2003, *L'analyse des musiques électroacoustiques : modèles et propositions*. Paris, L'Harmattan, 590 p.

Rubin, E., 1915, *Synsoplevede figurer*. Copenhague, Gyldendalske Boghandel. Édition allemande : *Visuell wahrgenommene Figuren. Studien in psychologischer Analyse* (Collett, P., trad.), 1921, 13 pl., 237 p.

Russolo, L., 1916, « L'art des bruits ». In Lista, G. (éd.), *L'art des bruits*. Lausanne, L'Âge d'Homme, 1975, 164 p.

Sacks, O., 1988, *L'homme qui prenait sa femme pour un chapeau*. Paris, Seuil, 318 p.

Saint-Loubry, B., 1997, *Perception de la qualité des archets de violons.* DEA. Atiam (LAM), université Paris 6.

Saldanha, E. L., & Corso, J., 1964, « Timbre Cues and the Identification of Musical Instruments ». *J. Acoust. Soc. Am.,* 36, p. 2021-2026.

Samuel, C., 1984, « Un entretien avec Pierre Boulez », in *Le Monde* daté du 14 novembre, cahier spécial « Forum des percussions », 14 novembre 1984-6 janvier 1985.

Sauveur, J., 1700-1713, *Collected writings on musical acoustics (Paris 1700-1713)*. Utrecht, The Diapason Press, 1984, 279 p.

Sauveur, J., 1701, *Principes d'acoustique et de musique ou système général des intervalles des sons* (édition de Paris). Mémoire de l'Académie royale des sciences. Genève, fac-similé Minkoff-Reprint 1973, 3 planches, 68 p.

Sauveur, J., 1702, « Application des sons harmoniques à la composition des jeux d'orgue ». In *Histoire de l'Académie royale des sciences – Année 1702*. Paris, Charles-Estienne Hochereau, p. 316-336.

Schaeffer, P., 1952, *À la recherche d'une musique concrète*. Paris, Seuil, 228 p.

Schaeffer, P., 1966, *Traité des objets musicaux*. Paris, Seuil, 712 p.

Schaeffer, P., 1985, « Varèse vingt ans après – Discussion suivant la communication de Jean Roy ». *La Revue musicale,* 383-385. Paris, Richard Masse Éditeur.

Schafer, R. M., 1979, *Le Paysage sonore* (Cail, O., trad.). Paris, JC Lattès, 391 p.

Schnupp, J., Nelken, I., & King, A., 2011, *Auditory Neuroscience – Making sense of sound* [démonstrations : www.auditoryneuroscience.com]. Cambridge (États-Unis), MIT, 356 p.

Schoenberg, A., 1911, *Traité d'harmonie* (Gubisch, G., trad.). Paris, JC Lattès, 1983, 519 p.

Schön, D., Gordon, R. L., & Besson, M., 2005, « Musical and linguistic processing in song perception ». *Annals of the New York Academy of Sciences,* 1060 (1), p. 71-81.

Schonle, J., & Horan, K., 1980, « The pitch of vibrato tones ». *J. Acoust. Soc. Am.,* 67, p. 246-252.

Schouten, J. F., 1962, *On the perception of sound and speech.* Conférence présentée à l'ICA (4th International Congress on Acoustics), Copenhague.

Schouten, J. F., 1968, *The perception of timbre.* Conférence présentée à l'ICA (6th International Congress on Acoustics), Tokyo.

Schouten, J. F., & 't Hart, J., 1965, « The strike note of bells ». *Neth. Acoust. Soc. Pub.,* 7, p. 8-19.

Schouten, J. F., Ritsma, R. J., & Lopes Cardozo, B., 1962, « Pitch of the Residue ». *J. Acoust. Soc. Am.,* 34 (8, part 2), p. 1418-1424.

Scotto Di Carlo, N., 2008, « Role of proprioceptive memory in a professional opera singer's absolute pitch. An experimental pilot study ». *Journal of Experimental Voice Research,* 1 (2), p. 34-39.

Seashore, C. E., 1938, *Psychology of Music*. New York, McGraw-Hill Book Company Inc., 408 p.

Sethares, W. A., 1998, *Tuning, Timbre, Spectrum, Scale*. Londres, Springer-Verlag, 345 p.

Shamma, S. A., & Micheyl, C., 2011, « Behind the scenes of Auditory Perception ». *Current Opinion in Neurobiology,* 20 (3), p. 361-366.

Shen, S., 1987, « Les anciennes cloches chinoises ». *Pour la Science,* juin, p. 100-108.

Shepard, R. N., 1964, « Circularity in judgments of relative pitch ». *J. Acoust. Soc. Am.,* 36, p. 2345-2353.

Shepard, R. N., 1992, *L'œil qui pense. Illusions, ambiguïtés et autres anomalies visuelles originales, avec un commentaire sur le jeu de l'esprit dans la perception et les beaux-arts* (Henry, J., trad.). Paris, Seuil, 234 p.

Slawson, A. W., 1968, « Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency ». *J. Acoust. Soc. Am.,* 43 (1), p. 87-101.

Slawson, W., 1985, *Sound Color.* Berkeley, University of California Press (livre + disque), 266 p.

Solomos, M., 2013, *De la musique au son ; l'émergence du son dans la musique des XXe-XXIe siècles*. Rennes, Presses universitaires de Rennes, 548 p.

Sotiropoulos, D., 2004, *Analyse acoustique et catégorisation d'un ensemble de qualités vocales pertinent pour la description de voix lyriques masculines.* Mémoire du DEA Atiam, Université Paris 6.

Souris, A., 1976, *Conditions de la musique et autres écrits*. Paris, ULB/CNRS, 312 p.

Stevens, P. S., 1978, *Les formes dans la nature* (Matricon, J., & Morello, D., trad.). Paris, Seuil, 230 p.

Stevens, S. S., & Davis, H., 1938, *Hearing – Its Psychology and Physiology*. New York, John Willey (Reprint, 1983, Acoustical Society of America/American Institute of Physics), 489 p.

Stockhausen, K., 1988, « ... Comment passe le temps... » (traduction de Ch. Meyer de « ... wie die Zeit vergeht... », article paru dans *Die Reihe*, n° 3, 1957). *Contrechamps,* 9, p. 26-65.

Strong, W., & Clark, M., 1966, « Perturbations of synthetic orchestral wind-instruments tones ». *J. Acoust. Soc. Am.,* 41 (2), p. 277-285.

Strong, W., & Clark, M., 1966, « Synthesis of wind-instrument tones ». *J. Acoust. Soc. Am.,* 41 (1), p. 39-52.

Sundberg, J., 1987, *The science of the singing voice*. Dekalb (Illinois), Northern Illinois University Press, 216 p.

Sundberg, J., 2001, « Level and center frequency of the singer's formant ». *J. of Voice,* 15 (2), p. 176-186.

Takeuchi, A. H., & Hulse, S. H., 1993, « Absolute Pitch ». *Psychological Bulletin,* 113 (2), p. 345-361.

Tamba, A., 1974, *La structure musicale du nô* (+ 2 disques). Paris, Klincksieck, 255 p.

Tarneaud, J., 1941, *Traité pratique de phonologie et de phoniatrie* (avec la collaboration de S. Borel-Maisonny). Paris, Librairie Maloine, 469 p.

Taylor, R., 1973, *Le bruit et ses méfaits* (Michel, W., & Olivier, C., trad.). Verviers, Gérard & Co, 285 p.

Thom, R., 1981, *Modèles mathématiques de la morphogenèse*. Paris, Christian Bourgois, 320 p.

Thomas, O., Monteil, M., & Touzé, C., 2011, « Idiophones à plaque et à coque : des cloches aux steelpans, acoustique des instruments à percussion ». *Musique et technique,* 6, p. 1771-3641.

Thompson, D. A., 1994, *Forme et croissance* (Teyssié, D., trad.), Paris, Seuil/CNRS, 336 p.

Titze, I. R., 1994, *Principles of voice production*. Englewood Cliffs (New Jersey), Prentice Hall, 354 p.

Tokumaru, Y., 1991, « Le timbre dans la musique japonaise ». In Barrière, J.- B. (éd.), *Le timbre – Métaphore pour la composition*. Paris, Christian Bourgois/IRCAM, p. 90-97.

Trân Van Khe, 1965, in « La vielle, le monocorde vietnamien, la cithare à 16 cordes ». *Bulletin du GAM (Groupe d'acoustique musicale),* n° 12. LAM, université Paris 6.

Trân Van Khe, 1968, « L'utilisation du sonagraphe dans l'étude du rythme », *Revue de Musicologie,* vol. LIV (2), p. 222-232.

Traube, C., & D'Alessandro, N., 2005, *Synthèse de l'imitation vocale et représentation graphique des gestes phonétiques sous-jacents à la description du timbre de la guitare classique*. Conférence présentée au CIM05 (Conference on Interdisciplinary Musicology), Montréal.

Trébinjac, S., 2008, *Le pouvoir en chantant* (vol. 2). Nanterre, Société d'ethnologie, 214 p.

Tyndall, J., 1869, *Le son* (Abbé Moigno, trad.). Paris, Gauthier-Villars, 381 p.

Vaissière, J., 2011, *La phonétique*. Paris, PUF, coll. Que sais-je ? (n° 637), 127 p.

Van Heuven, E. W., 1949, *Acoustical measurements on church-bells and carillons*. La Haye (Pays-Bas), Gebroeders Van Cleef, voir http://repository.tudelft.nl, 164 p.

Van Noorden, L. P., 1975, *Temporal coherence in the perception of tone sequences.* Thèse de doctorat, Technische Hogeschoel, Eindhoven [Document téléchargeable sur le site http://www.ipem.ugent.be/user/8].

Vandenheede, J., & Harvey, J., 1985, *Identity and Ambiguity: The Construction and Use of Timbral Transitions and Hybrids.* Ann Arbor, MI: MPublishing, University of Michigan Library, p. 97-102.

Vanderveer, N. J., 1979, *Ecological Acoustics: Human Perception of environmental sounds*. Thèse, Faculty of Cornell University. Ann Arbor (Michigan), University Microfilms n° 8004002, 240 p.

Vangenot, S., 2004, « L'énigme de l'oreille absolue ». In Nattiez, J.-J. (éd.), *Musiques – Une encyclopédie pour le xxe siècle*, vol. 2 *Les savoirs musicaux.* Paris, Actes Sud/Cité de la Musique, p. 527-539.

Varela, F., 1996, *Invitation aux sciences cognitives*. Paris, Seuil, coll. Points Sciences, 130 p.

Varela, F., Thompson, E., & Rosch, E., 1993, *L'inscription corporelle de l'esprit* (Havelange, V., trad.). Paris, Seuil, 380 p.

Varèse, E., 1983, *Écrits – Textes réunis et présentés par Louise Hirbour* (Léaud, C., trad.). Paris, Christian Bourgois, 217 p.

Vennard, W., 1967, *Singing, the mechanism and the technic.* New York, Carl Fischer, 275 p.

Virdung, S., 1511, *Musica Getutscht* (Meyer, C., trad.). Bâle, CNRS, éd. française, 1980, 157 p.

Vogel, C., 1999, *Étude sémiotique et acoustique de l'identification des signaux sonores d'avertissement en contexte urbain*. Thèse de l'université Paris 6.

Voort (van der), A. W. M., & Aarts, R. M., 2009, « Development of Dutch sound locators to detect airplanes (1927-1940) ». *NAG/DAGA*, p. 250-253.

Warren, R. W., 1994, « La perception des séquences acoustiques : intégration globale ou résolution temporelle » (Collart, M.-I., trad.). In McAdams, S. & Bigand, E. (éd.), *Penser les sons : psychologie cognitive de l'audition.* Paris, PUF, p. 41-73.

Warusfel, O., 2008, « La spatialisation du son ». *Pour la Science,* 373, p. 100-106.

Weil-Barais, A. (éd.), 1993, *L'homme cognitif*. Paris, PUF.

Weinreich, G., 1997, « Directional tone color ». *J. Acoust. Soc. Am.,* 101 (4), p. 2338-2346.

Wessel, D., 1979, « Low dimensional control of musical timbre », *Computer Music Journal*, 3 (2), p. 45-52.

Wessel, D., 1979, « Timbre space as a musical control struture ». *Computer Music Journal,* 3 (2), p. 45-52.

Winckel, F., 1960, *Vues nouvelles sur le monde des sons* (Moles, A., & Lequeux, J., trad.). Paris, Dunod, 144 p.

Witkowski, N., 2003, *Une histoire sentimentale des sciences*. Paris, Seuil, 331 p.

Wolfe, J., Garnier, M., & Smith, J., 2009, « Vocal tract resonances in speech, singing, and playing musical instruments ». *HFSP Journal* (http://hfsp.aip.org), 3 (1), p. 6-23.

Wu, Rung Shun, 1996. *Tradition et transformation : Le pasi but but, un chant polyphonique des Bunun de Taïwan*. Thèse, Université Paris X, Nanterre.

Wye, T., 2003, *Practice book for the flute*, vol. 4 *Intonation and vibrato*. Londres, Novello, p. 14.

Xenakis, I., 1963, « Musiques formelles ». *La Revue musicale,* 253-254. Paris, Richard Masse Éditeur, 232 p.

Yost, W. A., & Hafter, E. R., 1987, « Lateralization ». In Yost, W. A., & Gourevitch, G. (éd.), *Directionnal Hearing*. New York, Springer Verlag, p. 49-82.

Zatorre, R., 2003, « Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function ». *Nature Neuroscience,* 6 (7), p. 692-695.

Zatorre, R. J., Belin, P., & Penhune, V. B., 2002, « Structure and function of auditory cortex: music and speech ». *Trends in Cognitive Sciences,* 6 (1), p. 37-46.

Zwicker, E., & Fastl, H., 1999 (2e éd.), *Psychoacoustics, Facts and models*. Springer Verlag, 416 p.

4. Internet : sites et pages personnelles

Allan, R., 2012, Reed Organs in England – Enharmonic harmoniums. Consulté en 2014, http://tardis.dl.ac.uk/FreeReed/organ_book/node22.html

Anonyme, 2003, Panorama des musiques électroniques. Consulté en 2014, http://sonhors.free.fr/

Baskevitch, F., 2008, *Les représentations de la propagation du son, d'Aristote à l'Encyclopédie* (Thèse). Consulté en 2014, http://tel.archives-ouvertes.fr/tel-00423362/.

Benoit, S., Blouin, D., Dupont, J.-Y., & Emptoz, G., 2009, « Chronique d'une invention : le "phonautographe" d'Edouard-Léon Scott de Martinville (1817-1879) et les cercles parisiens de la science et de la technique ». *Documents pour l'histoire des techniques*, XVII. Consulté en 2014, http://dht.revues.org/5022.

Bettens, O., « Intonation juste à la Renaissance ». Consulté en 2014, www.virga.org/zarlino/

Bregman, A. S., 2015. Al Bregman Auditory Scene Analysis (ASA). Consulté en 2015. http://webpages.mcgill.ca/staff/Group2/abregm1/web/

Clerc, P.-A. « Discours sur la rhétorique musicale et plus particulièrement la rhétorique allemande entre 1600 et 1750 ». Consulté en 2014, téléchargeable en PDF à partir du titre de l'article.

Deutsch, D., 2013. Illusions and research. Consulté en 2015. http://deutsch.ucsd.edu/psychology/pages.php?i=201

Deutsch, D. http://deutsch.ucsd.edu/psychology/pages.php?i=101, page personnelle, et site http://philomel.com/musical_illusions, visité en août 2014.

Falstad, P., 2013, Acoustics and vibrations animations. Consulté en 2014. www.falstad.com/mathphysics.html

Fokker, A. D., 1955, « Equal temperament and the Thirty-one-keyed organ ». Consulté en 2014, www.huygens-fokker.org/docs/fokkerorg.html

Hai, T. Q., & Zemp, H., 2010, Analyses acoustique et physiologique des techniques vocales. Consulté en 2014, http://crem-cnrs.fr/clefs-ecoute/animations/diphonique/hai1.html

Henrich, N., 2012, « Physiologie de la voix chantée : vibrations laryngées et adaptations phono-résonantielles ». *40es entretiens de médecine physique et de réadaptation*. Consulté en 2014. http://hal.archives-ouvertes.fr/docs/00/68/06/92/PDF/Henrich_EMPM_2012_physiovoixchantee.pdf.

Labussière, A., 2008, « D'un pentatonique l'autre : processus, structure, gestuelle, polarités dans le chant traditionnel à voix nue ». *Musimédiane*, n°3 (musiques non écrites). Consulté en 2014, www.musimediane.com.

Museum Waalsdorp, Dispositifs d'aide au repérage aérien, années 1915-1935. Consulté en 2015. http://www.museumwaalsdorp.nl/en/airacous.html

Patterson, R. D. Page personnelle. Consultée en 2014. www.neuroscience.cam.ac.uk/directory/profile.php?rdp1

Pels, I. B. J. A., 1950, « L'orgue accordé en cinquièmes de ton du musée Teyler à Haarlem ». Consulté en 2014. www.huygens-fokker.org/docs/pels.html

Pujol, R., 2007, Promenade autour de la cochlée. Consultée en 2014, www.cochlee.org

Russel, D., 2013, Démonstrations animées. Consultées en 2014, www.gmi.edu/~drussell/Demos.html

Shimojo, S., & collaborators, 1997. Hall of illusions. Consulté en 2013, www.psychologie.tu-dresden.de/i1/kaw/diverses%20Material/www.illusionworks.com/html/hall_of_illusions.html

Terhardt, E., 2007, Page personnelle. Consultée en 2014. www.mmk.e-technik.tumuenchen.de/persons/ter.html

Trân Quang Hai. Page personnelle. Consultée en 2015. site : http://tranquanghai.info/index.php?lang=fr

Wertheimer, M., 1880-1943, voir : page Wikipedia en allemand. Consultée en 2013. http://de.wikipedia.org/wiki/Max_Wertheimer

Wolfe, J., 2012, Physclips – Waves and sound. Consulté en 2014. www.animations.physics.unsw.edu.au/waves-sound/

Wolfgang, C., & Rousseau, J.-J., 2013, Ondes progressives et ondes stationnaires – Simulations dynamiques interactives. Consultées en 2014. http://webapps.fund p.ac.be/didactique/ondes/index.php

5. Logiciels

Wavesurfer – Logiciel open source. Copyright © 2000-2011 Jonas Beskow et Kare Sjolander. Développé au Centre for Speech Technology (CTT) at KTH, Stockholm, Suède. Manuel en ligne : www.speech.kth.se/wavesurfer/man.html

Praat – Logiciel libre. Copyright © 1992-2011 Paul Boersma et David Weenink. Développé à Phonetic Sciences, Université d'Amsterdam Spuistraat 210 – 1012VT Amsterdam – Pays-Bas. Téléchargement du logiciel et du manuel : www.praat.org

Audiosculpt (version 2.9) – Copyright © 1995-2010 IRCAM Voir site du forum : http://forumnet.ircam.fr

D'utiles indications sur les logiciels d'accès libres sont disponibles sur le site http://en.wikipedia.org/wiki/List_of_free_software_for_audio

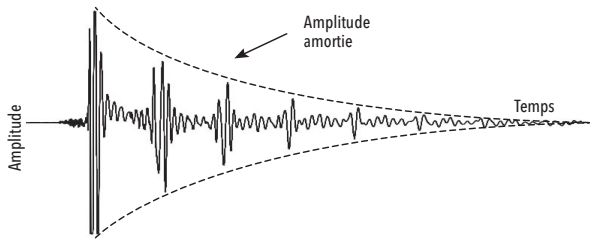
ANNEXE G

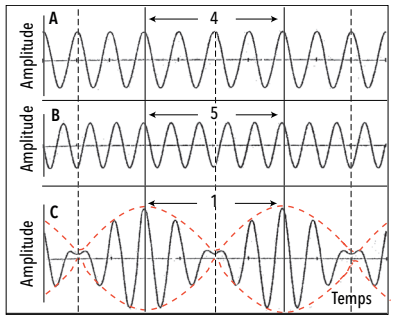
Glossaire

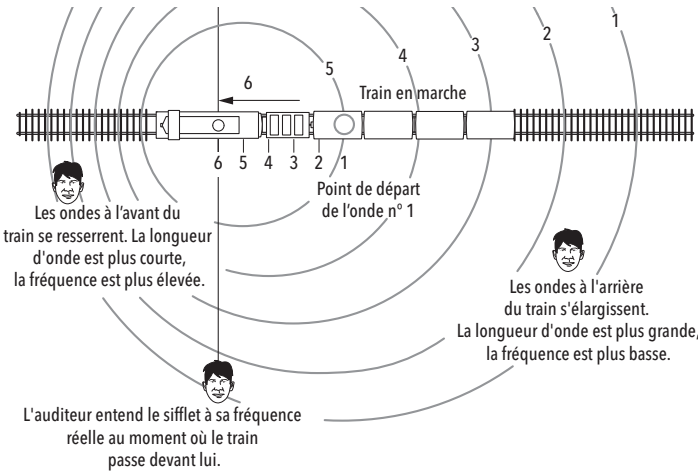
Le glossaire offre les définitions de termes en usage dans ce livre en insistant sur le fait que certains d'entre eux prennent des sens différents selon qu'il s'agit de musique ou d'acoustique. Les notions spécifiques à l'acoustique complètent et précisent celles qui sont présentées dans les chapitres 1 et 2. Le lecteur pourra les approfondir en se reportant aux ouvrages spécialisés. Pour la compréhension des modes vibratoires, nous recommandons vivement la consultation des sites d'animation. Les livres et les sites, qu'on trouvera dans la bibliographie (annexe E), sont référencés comme suit dans le tableau :

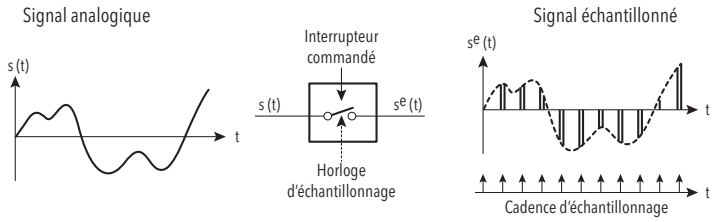
- livres : **CK** (Chaigne & Kergomard) ; **EL** (E. Leipp, 2011) ; **TR** (T. Rossing) ; **TS** (Techniques du son).
- sites d'animation : ***sJW*** (J. Wolfe) ; ***sDR*** (D. Russel) ; ***sPF*** (P. Falstad) ; ***sWR*** (Wolfgang & Rousseau).

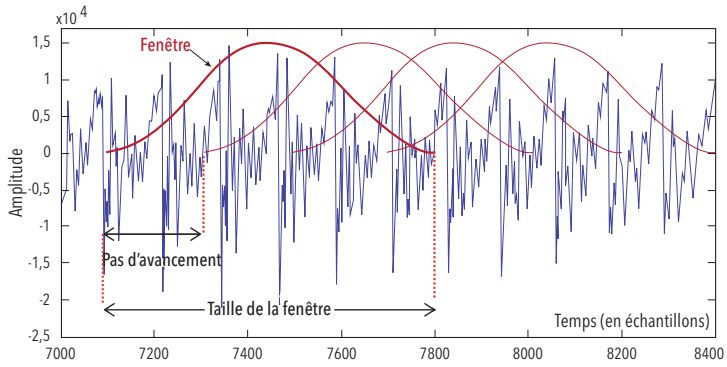
Remarque : les notes de bas de page sont rassemblées à la fin du tableau, page 530.

TERMES	DÉFINITIONS ET EXEMPLES
Amortissement Figure G.1	 Grandeur caractérisant la dissipation d'énergie dans un système mécanique. Les instruments en oscillations libres vibrent selon des mouvements périodiques d'amplitude globalement décroissante : l'amplitude diminue au cours du temps, mais la fréquence vibratoire reste constante. Le taux d'amortissement, qui dépend des effets des frottements du système vibrant ainsi que de la nature des matériaux, a une incidence sur la largeur des pics de résonance. Soient deux plaques suspendues de mêmes dimensions que l'on excite par impulsion, l'une en bois, l'autre en métal. La plaque en bois produit un son bref, de hauteur complexe. La plaque en métal produit un son notablement plus long, dans lequel plusieurs notes sont repérables. L'amortissement des ondes est plus faible dans le métal que dans le bois. Voir chapitre 7, § 2.2.1.
Amplitude	L'amplitude d'une déformation, d'une vibration est l'écart maximal (en m) par rapport à la position initiale (repos).

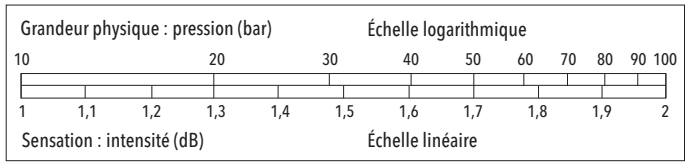
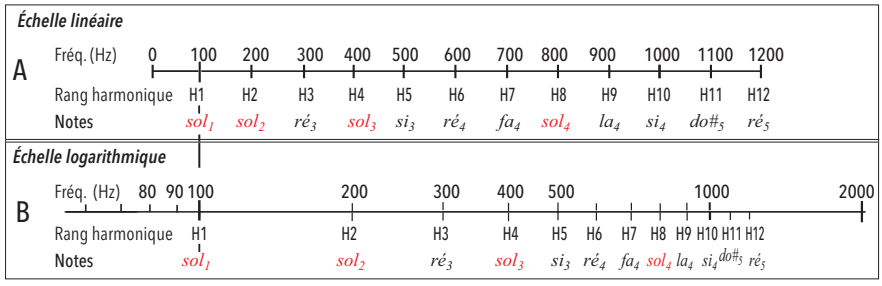
TERMES	DÉFINITIONS ET EXEMPLES
Battement Figure G.2	 Les battements résultent de la combinaison des amplitudes de deux composantes de fréquences *f* voisines. Soient par exemple (figure G.2) deux sons purs A et B avec *f*A = 4Hz et *f*B = 5 Hz. Leurs amplitudes s'additionnent en produisant une onde résultante C dont l'enveloppe est modulée en amplitude à la fréquence du battement qui est *f*B – *f*A = 1 Hz. Le nombre de battements par seconde correspond donc très exactement à la différence de fréquence, en Hz, entre deux fréquences voisines. Si nous considérons l'intervalle musical, le nombre de battements qui le caractérise change avec la tessiture. En effet, quand la différence entre les deux fréquences croît, l'incidence perceptive change de nature, passant de la perception d'une modulation d'amplitude à celle de rugosité, puis aux sons différentiels. Voir chapitre 3, § 3.2.5. Les battements jouent un rôle important dans l'accordage des instruments de musique. Voir chapitre 8, § 2.
Bottom-up	Expression anglaise, littéralement « du bas vers le haut ». En perception sonore, cette expression figure la succession des opérations concernant l'audition, depuis l'oreille jusqu'aux centres nerveux supérieurs. Voir chapitre 4, § 1.2.
Bruit	Terme polysémique. Dans cet ouvrage le mot *bruit* désigne des phénomènes chaotiques, sans connotation hédonique : bruits de souffle de la flûte, de la voix chuchotée, bruit de fond d'un enregistrement. Famille de sons dont le contenu spectral est quasi continu. Le *bruit blanc* contient toutes les fréquences du spectre apparaissant de façon aléatoire, avec la même énergie.
Célérité	Terme désignant la vitesse de propagation d'une onde dans un milieu donné. La célérité (c) dépend de la densité, de l'homogénéité du milieu, et de la température[1], surtout pour les gaz. La célérité du son dans l'air varie de 331m/s à 0 °C à 349 m/s à 30 °C, ce qui a une incidence directe sur le diapason des instruments à vent. Elle est beaucoup plus élevée dans l'eau (environ 1435 m/s) et dans l'acier (environ 5000 m/s). Dans le bois, qui est un matériau anisotrope, la célérité varie selon que la propagation a lieu dans le sens des fibres (5000 m/s pour l'épicéa) ou dans le sens transversal aux fibres (1000 m/s pour l'épicéa, à 15 % d'humidité).
Discret	Un signal discret est constitué d'échantillons séparés (cas numérique). Il s'oppose au signal continu (cas analogique). Par extension, un spectre harmonique peut être qualifié de discret, par opposition au spectre continu d'un bruit blanc. La numérisation du son, ou discrétisation, consiste à convertir une courbe continue en une suite finie d'échantillons prélevés à intervalles de temps égaux (voir *Échantillonnage*).

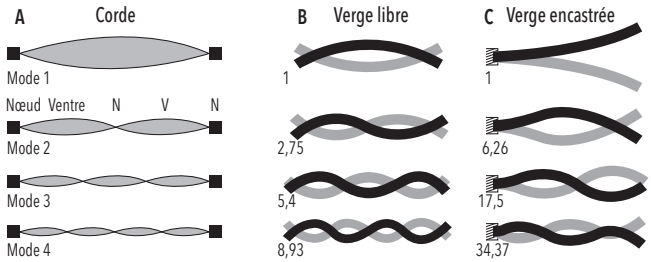
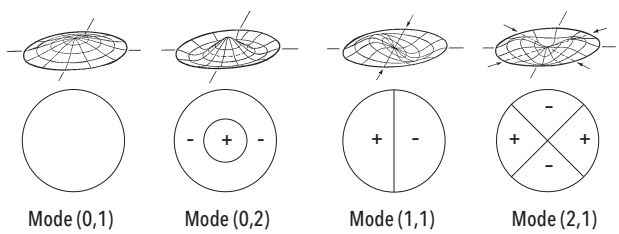
TERMES	DÉFINITIONS ET EXEMPLES
Doppler (effet)	 Figure G.3 Voir site ***sJW***. D'après *Science Club*, n° 26, avril 1966, p.12. L'effet Doppler se produit lorsqu'une source vibratoire est en mouvement par rapport au récepteur. La combinaison de la vitesse de propagation de l'onde avec celle du déplacement de la source a pour conséquence de générer des longueurs d'onde différentes à l'avant et à l'arrière de la source. C'est la seule situation pour laquelle la longueur d'onde (λ) change durant la propagation dans un milieu donné. Par voie de conséquence, la fréquence du son entendu change aussi. On a : $\lambda = (c \pm v)/f$ où (v) est la vitesse (m/s) du véhicule. *Exemple numérique*. Une voiture klaxonne à 400 Hz en se déplaçant à 72 km/h. Quelle est, pour un auditeur immobile, la variation de fréquence entre le son capté avant son passage et celui qui se propage après son passage ? Le véhicule se déplace à 20m/s. Pour $\lambda = 340$ m/s, la vitesse relative de la source varie de 360 à 320 m/s, donc dans un rapport de 9/8 (1,125), ce qui correspond à un intervalle d'un ton.
Dynamique (en acoustique)	Ce terme a ici pour sens général l'écart (en dB) entre les sons les plus faibles et les sons les plus forts. On peut définir la dynamique pour une note, pour l'ensemble des sons d'un instrument, pour un groupe d'instruments. Voir le chapitre 5 § 4. La limite supérieure augmente lorsque plusieurs instruments sont combinés (orchestre, registres d'un orgue), mais la limite inférieure ne peut descendre en dessous du bruit de fond qui est de l'ordre de 40 dB dans un lieu calme (salle de concert). Pour un *tutti* d'orchestre à 110 dB, la dynamique est donc de l'ordre de 70 dB. L'enregistrement numérique permet aujourd'hui de capter et de reproduire de grandes dynamiques. Cependant, l'écoute dans différents lieux de vie, en particulier les transports (train, métro, avion), conduisent souvent à limiter la dynamique des enregistrements en comprimant le signal, faute de quoi les sons les plus faibles sont noyés dans le bruit de fond.

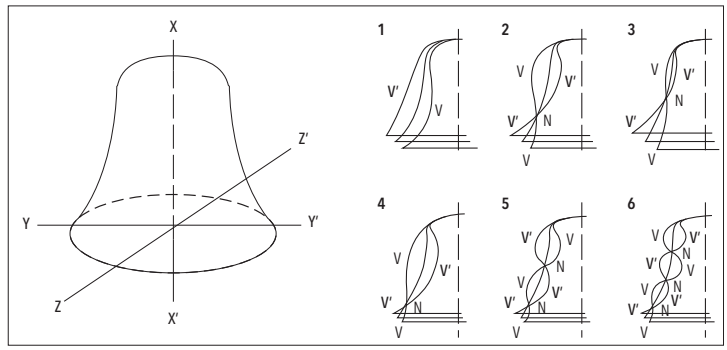
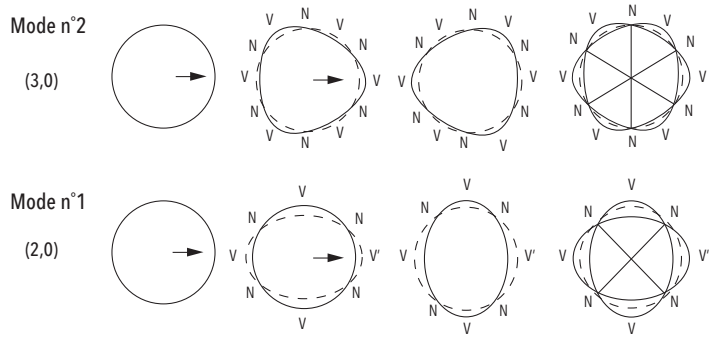
TERMES	DÉFINITIONS ET EXEMPLES
Échantillonnage Fréquence d'échantillonnage (*fe*) ; nombre de bits Figure G.4 D'après Fromentel, A., 2010, p. 331.	 L'échantillonnage consiste à transformer un signal analogique (continu) en signal numérique (discret), en prélevant les valeurs des coordonnées de plusieurs points définis à intervalles de temps égaux. Un *échantillon sonore* (ou point) est défini par les valeurs de l'amplitude et du temps qu'il a sur la courbe analogique (voir **TS**, chapitre 8). Combien de points sont nécessaires pour numériser une courbe donnée ? Le *théorème de Shannon* stipule qu'il faut au minimum 2 points par période pour représenter une fréquence sinusoïdale. Il faudra donc 880 points pour un son pur de 440 Hz (la_3 du diapason). En pratique, ce nombre est légèrement supérieur, c'est pourquoi la fréquence d'échantillonnage *fe*, qui doit garantir la bonne restitution de la fréquence la plus élevée d'un signal sonore, est toujours un peu supérieure au double de la limite supérieure de la bande passante. Ainsi la fréquence d'échantillonnage du CompactDisc, qui est de 44,1 kHz, permet de couvrir la bande audible humaine dont le maximum est de 20 kHz. L'amplitude du signal (échelle verticale) est codée numériquement en *bits*. Le format du nombre de valeurs en bits (valeurs binaires) rend compte de la finesse de *quantification* du signal, donc de la dynamique. On a : 4 bits = 2^4 = 16 valeurs ; 8 bits = 2^8 = 256 valeurs ; 16 bits = 2^{16} = 65 536 valeurs par échantillon. À titre d'exemple, une seconde de son stéréophonique (2 pistes) échantillonnée à 44,1 kHz sur 16 bits produit un fichier numérique d'environ 173 ko^2, soit un peu plus de 10 Mo pour une minute, et environ 622 Mo pour une heure.
Élasticité	Propriété d'un matériau capable de reprendre sa forme d'origine lorsqu'il a été déformé. Un matériau est élastique lorsque sa structure est telle qu'il oppose à une force qui le déforme une force contraire dite force de rappel, qui tend à ramener la zone déformée à sa position de départ. Une corde élastique revient à sa longueur initiale lorsqu'elle a été tendue.
Filtrage	Opération qui consiste à supprimer certaines composantes d'un son. Un filtre *passe-haut* élimine les fréquences inférieures à sa limite, dénommée fréquence de coupure (en Hz) ; à l'inverse, un filtre *passe-bas* élimine les fréquences supérieures à cette limite. Un filtre *passe-bande* ne laisse passer que les fréquences comprises entre deux limites de fréquence. S'il est très étroit, il peut ne laisser passer qu'un harmonique. Voir le chapitre 1, § 4.
Fondamental	Ce terme a différents sens. En musique, c'est la note fondamentale d'un accord. En facture, c'est le son le plus grave d'un tuyau, ou d'une corde. En acoustique, c'est la fréquence fondamentale d'un son. Le *son fondamental* d'une série harmonique est l'harmonique n° 1. Le son fondamental d'un tuyau, par exemple d'une flûte, est la fréquence du premier mode vibratoire, que l'on nomme aussi le partiel 1 du tuyau (voir *Partiel*).
Formant	La notion de formant a été développée en acoustique de la voix, mais on l'emploie aussi pour caractériser le spectre de certains instruments de musique. Un formant correspond à un maximum de la courbe spectrale, dans lequel une ou plusieurs composantes sont renforcées. Il traduit la présence de résonances dans la structure vibrante. Un formant est caractérisé par sa *fréquence centrale* et par sa *largeur de bande*. Celle-ci est définie comme la largeur de l'enveloppe spectrale (ou intervalle en Hz) prise 3 dB au-dessous du maximum. Dans le domaine de la voix, les formants correspondent aux principaux modes de vibration du conduit vocal. Les trois premiers, notés F1, F2 et F3, caractérisent les voyelles. Voir le chapitre 7, § 2.3.5 pour les instruments et le chapitre 9, § 1.4.2 pour la voix.

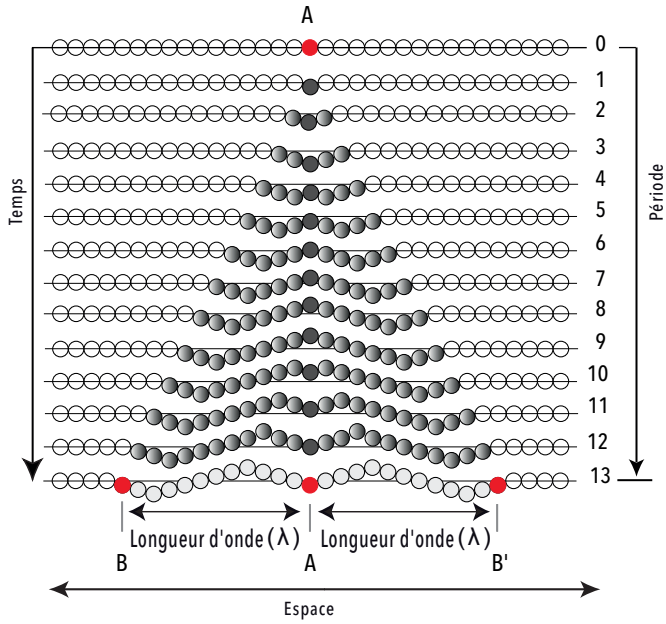
TERMES	DÉFINITIONS ET EXEMPLES
Fourier (Théorème de)	Le théorème de Fourier stipule que tout son périodique de fréquence *f* est décomposable en une série de sons purs dont les fréquences sont des multiples entiers de *f*. Le mouvement périodique le plus simple, sinusoïdal, est une fréquence pure (exemple, la fréquence propre d'un mode vibratoire). Les mouvements périodiques complexes des sons musicaux comprennent un grand nombre de composantes de fréquences différentes, toutes multiples de la plus grave : ce sont les harmoniques. On peut décomposer la forme d'onde d'un son périodique complexe en ondes de fréquences pures : c'est l'analyse. Inversement, on peut construire par synthèse la forme d'onde d'un son périodique complexe par addition des amplitudes des composantes séparées, en respectant leurs phases relatives. Voir *Phase*.
Fourier (Transformée de)	 Figure G.5 @ Alain Lithaud, Ircam - Centre Pompidou. Le théorème de Fourier ne s'applique en toute rigueur qu'à un signal parfaitement périodique, ce qui est un concept purement théorique. Pour les signaux réels, de durée limitée, on utilise la transformée de Fourier (TF), qui permet de décomposer un signal quelconque en une somme d'un grand nombre de sinusoïdes voisines et non plus multiples du fondamental. En pratique, on a besoin de connaître, non pas le spectre de l'ensemble du signal, mais sa valeur instantanée et ses variations. Ceci est obtenu en découpant le signal en tranches de temps successives, au moyen d'une opération appelée *fenêtrage* : c'est la Transformée de Fourier *glissante*, ou à *Court Terme* (TFCT). La taille de la fenêtre est donnée en durée (secondes ou millisecondes) ou en nombre d'échantillons[3], ce qui revient au même si l'on connaît la fréquence d'échantillonnage. Pour éviter des perturbations aux extrémités de la fenêtre, cette dernière – par exemple la fenêtre de Hanning – possède une forme en arche de sinusoïde. La durée et la forme de la fenêtre ont une grande incidence sur le spectre obtenu : une fenêtre large entraîne une bonne résolution fréquentielle et une faible résolution temporelle, alors qu'une fenêtre étroite produit l'effet inverse. L'intervalle entre deux fenêtres successives est appelé *pas d'avancement*. Pour ne pas perdre d'information avec une fenêtre de Hanning, le pas d'avancement doit être inférieur à la moitié de la largeur de la fenêtre. La TFCT est une opération coûteuse en temps de calcul. L'invention d'un algorithme très rapide dans les années 1960, connu sous le nom de *Fast Fourier Transform* (FFT) et reposant sur une décomposition binaire des étapes de calcul, en a permis la généralisation à l'analyse de tous les signaux, et en particulier aux signaux audio.
Fréquence	La fréquence *f* en Hz d'un phénomène périodique est le nombre de cycles accomplis par unité de temps (seconde). La fréquence *f* est inverse de la période T.

TERMES	DÉFINITIONS ET EXEMPLES
Harmonique (son)	Le son *harmonique* (mot masculin ou féminin) a un sens précis en acoustique, distinct de ceux qu'il prend dans la pratique musicale (voir aussi *Harmonique effleuré*). 1/ Un son *harmonique* est synonyme de son périodique. Seuls les sons entretenus sont composés d'harmoniques. Voir chapitre 1, § 5. 2/ Un *harmonique* isolé (d'une série harmonique) est un son pur. On peut mettre en évidence un harmonique d'un son complexe en le sélectionnant par résonance (ou par filtrage).Voir chapitre 1, § 4. La technique de renforcement des harmoniques par les résonances de la cavité buccale permet de jouer des mélodies spectrales ou harmoniques, comme le chant *diphonique* et par extension (car ce ne sont pas des sons entretenus) l'arc musical et la guimbarde. Voir chapitre 1, § 6 ; chapitre 6, § 2.7.3 ; chapitre 8, § 1.3.
Harmonique (série)	Selon le théorème de Fourier, un mouvement périodique complexe de période T (ou de fréquence *f*) peut être décrit comme une somme de mouvements élémentaires de périodes T, T/2, T/3 etc. (ou de fréquences *f*, 2*f*, 3*f*, etc.), avec les amplitudes et les phases appropriées. Les intervalles entre les sons d'une série harmonique, définis par le rapport de leurs fréquences, sont immuables. Ils forment une suite décroissante d'intervalles bien connue en musique, qui est, pour les 5 premiers harmoniques : octave, quinte, quarte, tierce majeure, tierce mineure. Voir chapitre 1, § 5.2.2.
Harmonique effleuré	Jouer un harmonique *effleuré* sur un instrument à cordes, c'est solliciter un nouveau fondamental correspondant à un mode propre de la corde, autre que celui généralement utilisé, qui est le mode le plus grave ou fondamental. Lorsqu'il est entretenu, un harmonique effleuré possède lui-même plusieurs harmoniques. Il serait donc plus exact de parler de « partiel » de la corde. Cette distinction est importante, car les harmoniques (d'un son) sont individuellement des fréquences pures. Voir chapitre 1, § 5.2.2.
Harmoniques d'un tuyau	De façon similaire aux harmoniques effleurés joués sur une corde, les instrumentistes produisent les « harmoniques » d'un tuyau en augmentant la pression d'alimentation ou par le moyen de doigtés appropriés. Il s'agit aussi des « partiels » des différents modes vibratoires du tuyau. Voir chapitre 1, § 5.2.2, les partiels d'un tuba ; chapitre 1, § 6.2, les partiels d'une flûte roumaine (tilinca) ; chapitre 2, § 4.4 et § 4.5, les partiels d'une flûte traversière et d'une clarinette.
Harmoniques (écoute)	Il faut des conditions particulières pour entendre isolément les harmoniques d'un son ; voir chapitre 1, § 4.3. L'oreille est avant tout sensible à la fréquence fondamentale et à la couleur globale du son (sonorité) qui est en rapport avec le contenu harmonique.
Inharmonique	Se dit d'un son ou d'une vibration qui ne sont pas périodiques. Un son non entretenu comporte toujours une certaine quantité d'inharmonicité due à la raideur du matériau vibrant. La 4e corde de la guitare (*sol*) est plus inharmonique que les autres, et partant, plus difficile à accorder. Le *coefficient d'inharmonicité* exprime l'écart en fréquence des partiels par rapport aux fréquences des harmoniques de même rang. L'inharmonicité augmente avec le rang de la composante. Elle confère une sonorité particulière aux sons du piano, notammnent dans les deux octaves basses, là où l'oreille est sensible aux partiels de rang élevé (20 à 30) qui sont les plus inharmoniques. Voir chapitre 6, § 3.
Intensité	Terme dont le sens est multiple. L'intensité, qui est liée à l'amplitude de la vibration sonore, correspond en première approximation à l'impression qu'un son est plus ou moins fort. Mesurer l'intensité, c'est mesurer l'énergie – ou travail fourni par l'onde sonore – qui passe à travers une surface unitaire, en un temps donné. Cette mesure est difficile. En pratique, on utilise un appareil spécialisé, le *décibelmètre*, muni d'un microphone étalonné, calibré pour une pression de référence. **Intensité absolue**. La *pression de référence,* qui est de 2.10^{-5} pascal (ou newton/m^2), est définie comme le seuil d'audibilité statistique pour la fréquence 1000 Hz. Elle correspond à un mouvement vibratoire extrêmement faible, d'amplitude 10^{-11} m, et à une puissance de 10^{-12} watt/m^2. Le « 0 dB » n'est pas perceptible puisqu'il existe toujours un bruit de fond environnant, même dans une chambre anéchoïque. **Intensité relative**. La plupart du temps, nous manipulons des valeurs relatives de l'intensité : le *rapport d'intensité* entre deux sources, entre deux composantes d'un son. Voir chapitre 3, § 2, et tout le chapitre 5.

TERMES	DÉFINITIONS ET EXEMPLES
Intervalle (perception)	Écart mesuré par le rapport entre deux valeurs d'un paramètre dimensionnel : fréquence, intensité, durée. Les intervalles musicaux correspondent aux rapports de fréquences entre deux sons successifs ou simultanés. L'interprétation des intervalles musicaux sur une analyse acoustique dépend du type d'échelle des fréquences : linéaire ou logarithmique. Voir chapitre 2, § 3.7.
Interférences	Lorsque deux ondes aériennes périodiques se croisent en un point donné de l'espace, leurs amplitudes se combinent. L'état local du milieu dépend de plusieurs facteurs : la fréquence, l'amplitude de chaque onde et leurs phases relatives. Voir *Phase* et le site ***sJW***.
Linéaire, Non linéaire	On dit d'un système vibrant qu'il est *linéaire* dans une zone donnée de son fonctionnement, lorsque soumis à un accroissement donné de l'excitation il y répond de façon proportionnelle, sans apporter d'élément nouveau. La plupart des instruments de musique mécaniques, et dans une certaine mesure l'oreille (voir chapitre 3, § 1.5.4 et § 2.2), se comportent de façon *non linéaire*, puisque l'augmentation d'amplitude de l'excitation a pour effet d'augmenter le nombre d'harmoniques dans le son produit. Voir le chapitre 5, § 3.1 : le *crescendo* musical. Les sons résultants sont des produits non linéaires de la combinaison de deux sources. Voir le chapitre 8, § 2.3. Les systèmes vibrants non linéaires relèvent de techniques mathématiques spécifiques et sont difficiles à modéliser. Il y a toujours avantage, lorsque c'est possible, à *linéariser* un modèle.
Logarithmique (échelle)	On représente généralement la variation des grandeurs physiques (croissance ou décroissance) sur des axes gradués. Si l'accroissement est de type additif, c'est-à-dire qu'il résulte de l'ajout de quantités égales, on représente les variations de la grandeur sur une *échelle linéaire*. Si l'accroissement est de type multiplicatif, c'est-à-dire qu'il correspond à une multiplication de la grandeur par un nombre constant, par exemple un facteur 10, on utilise un *échelle logarithmique* (à base 10). Une autre façon de procéder consiste à changer d'unité de mesure. En adoptant des unités logarithmiques comme le dB ou le cent, il est possible de travailler sur une graduation linéaire. Le choix de l'un ou de l'autre dépend évidemment des problèmes à traiter. Voici deux exemples. **La loi de Fechner-Weber**  Figure G.6 La figure G.6 représente la correspondance entre les variations d'une grandeur physique, la pression acoustique mesurée en bar, et l'intensité exprimée en dB, mieux représentative de la sensation. Voir chapitre 3, § 2.4 **Les fréquences des harmoniques et les intervalles musicaux**  Figure G.7 Comparez figure G.7 les deux représentations graphiques de la suite des harmoniques d'un son. En A, l'échelle linéaire rend compte de l'équidistance (en Hz) des harmoniques. En B, l'échelle logarithmique rend compte de la grandeur perceptive des intervalles musicaux, régulièrement décroissants. Voir chapitre 2, § 3.7.

TERMES	DÉFINITIONS ET EXEMPLES
Longueur d'onde	La longueur d'onde λ ou *périodicité spatiale* est la distance (en m) entre 2 points successifs en phase, dont les mouvements s'effectuent de façon synchrone. Pour un milieu donné, la longueur d'onde croît avec la célérité (c) et est inversement proportionnelle à la fréquence. $\lambda = c/f$. Exemples : dans l'air à 15 °C, λ = 3,40 m pour un son de 100 Hz (environ sol_1) ; λ = 0,34 m pour un son de 1000 Hz (environ si_4). La longueur d'onde varie avec la densité du milieu de propagation. Voir le phénomène dit de la voix « hélium », chapitre 7, § 2.4.1.
Mode vibratoire (mode propre)	Les modes vibratoires d'un système – corde, colonne d'air, plaque – sont caractérisés individuellement par une fréquence et une déformée (voir chapitre 1, § 2.4). Un *nœud* (N) est un point (ou une ligne de points) immobile, de part et d'autre duquel se produisent les mouvements alternés des *zones ventrales* (V). Leur étude relève de la dynamique des solides et de l'aérodynamique, et fait l'objet d'un formalisme mathématique très complexe, même lorsqu'il s'agit de corps de géométrie régulière, construits dans un matériau homogène dans toutes leurs dimensions. Voir ci-dessous, quelques exemples schématiques de modes vibratoires.
Modes 1D Figure G.8 Voir les animations sur les sites ***sJW*** ; ***sDR*** ; ***sPF***.	 **Corps à une dimension (1D) – Corde, verge** Exemples de modes vibratoires de structures à une dimension : **A** : corde vibrante. **B** : verge libre aux deux bouts (barre de vibraphone). **C** : verge encastrée à un bout et libre à l'autre (diapason, orgue de cristal). Les valeurs numériques indiquent les rapports de fréquences d'un mode *n* par comparaison au mode le plus grave, *n* = 1.
Modes 2D Figure G.9 Voir les animations des sites ***sJW***, ***sDR***, ***sPF***.	 **Corps à deux dimensions (2D) – Plaque mince** Premiers modes vibratoires d'une membrane circulaire fixée sur ses bords (appui ou encastrement). Il faut deux nombres pour définir un mode 2D : (*m*) le nombre de diamètres nodaux, et (*n*) le nombre de cercles nodaux. La plaque étant fixée sur son pourtour, tous les modes ont un cercle nodal. Pour chacun des quatre modes dessinés, une vue en perspective est placée en vis-à-vis du schéma. Les zones de même signe (+ ou –) sont en phase.

TERMES	DÉFINITIONS ET EXEMPLES
Modes de coque (axisymétrique)	 Figure G.10 Voir les animations des sites ***sJW***, ***sDR***, ***sPF***. D'après Lerhr, A., 1976, figure 13.  Figure G.11 Ibid. D'après Lerhr, A., 1976, figure 14. **Coques de type cloche, verre** Les modes vibratoires d'une cloche, d'un verre, doivent être explorés dans les trois dimensions (X, Y, Z), mais l'objet étant symétrique par rapport à l'axe vertical, deux plans suffisent. En haut, déformation du profil de la cloche dans le plan XY. En bas, déformations du bord circulaire de la cloche dans le plan YZ. Deux modes sont représentés selon quatre étapes temporelles (de gauche à droite). V et N désignent les Nœuds et Ventres de vibration. Définition de *coque* : voir **CK**, p. 36.
Modulation	La modulation est, dans un sens très général, une variation temporelle régulière (périodique) appliquée à un signal le plus souvent complexe. La modulation peut s'appliquer à l'amplitude ou à la fréquence du son. Utilisées pour le codage et la transmission de signaux sonores et plus récemment pour la synthèse sonore, les modulations sont maintenant un domaine technologique à part entière, doté d'une théorie mathématique spécifique. Mais bien avant les développements de l'électroacoustique, les modulations ont été exploitées en musique où les exemples abondent, en particulier les phénomènes désignés sous le nom général de « vibrato ». Voir chapitre 3, § 3.2 ; chapitre 6, § 4 ; chapitre 8, § 2.4.
Onde	Une onde est la propagation d'une perturbation produisant sur son passage une variation réversible de l'état local du milieu. Une onde transporte de l'énergie sans transporter de matière. Site ***: sJW***.

TERMES	DÉFINITIONS ET EXEMPLES
Onde, propagation Figure G.12	 La figure représente, de bas en haut, les différents états de la surface d'un plan d'eau sur lequel se propage une onde. Le temps est figuré sur l'axe vertical. Au temps 0, la surface est plane. L'impact central produit un mouvement transversal de la surface, initialement localisé au point A (temps 1). Dans les instants suivant, les particules d'eau (figurées par des boules) oscillent localement dans l'axe vertical, en reproduisant le mouvement initial avec un retard temporel (décalage de phase) dû au temps de propagation et l'onde de mouvement se propage ainsi de proche en proche aux particules de plus en plus éloignées du point d'impact. Au bout d'un certain temps (ligne 13), deux boules, B et B' démarrent leur oscillation verticale, en synchronisme avec la boule A qui amorce une nouvelle oscillation : les trois boules sont *en phase*. Le phénomène se produit au terme d'*une périod*e. Les boules B, A, d'une part, et A, B', d'autre part, sont séparées dans l'espace par une *longueur d'onde*. Dans cet exemple, nous supposons que chaque boule oscille verticalement par rapport à sa position d'équilibre et que les oscillations se reproduisent de façon régulière. Site : ***sWR***.

TERMES	DÉFINITIONS ET EXEMPLES
Onde (front)	

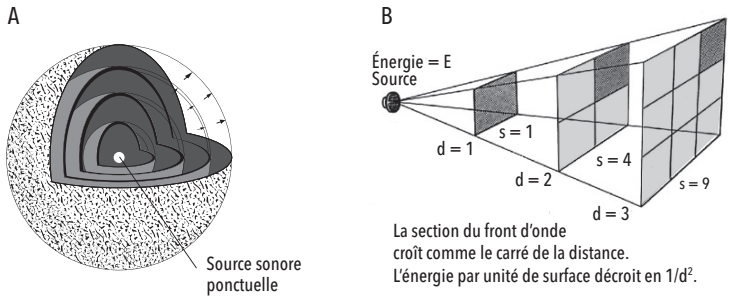


Figure G.13

D'après *Revue Science Club*, n° 26, avril 1966, p. 9.

À l'air libre, une onde se propage sous forme de sphères concentriques de dimensions croissantes (figure G.13-A). L'énergie initiale se trouve donc répartie sur une surface qui croît très rapidement, et l'intensité sonore que l'on peut capter en un point décroît en raison inverse du carré de la distance à la source (figure G.13-B). Lorsque l'onde se propage dans un conduit (tuyau, canalisation), la surface du *front d'onde* reste constante. L'intensité perçue à distance est étonnamment conservée.

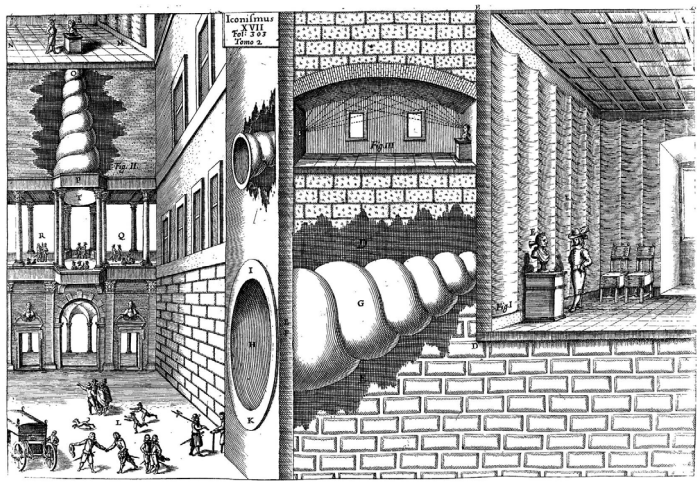
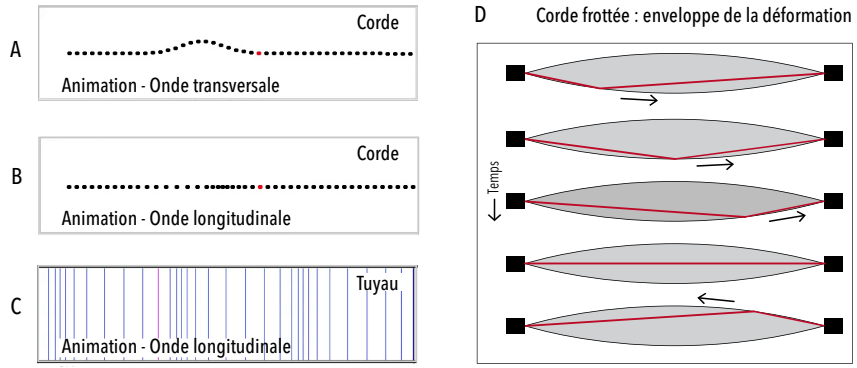
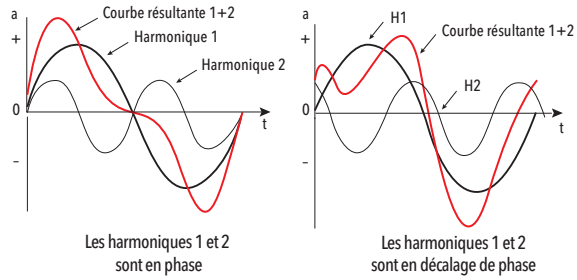


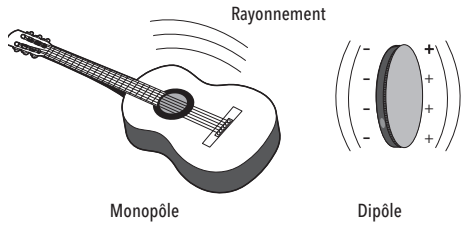
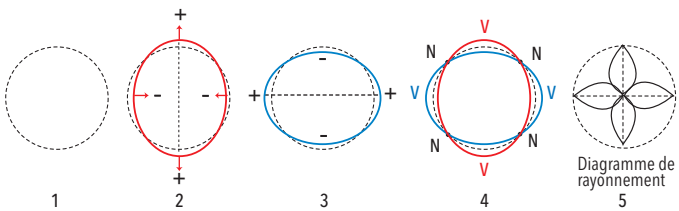
Figure G.14

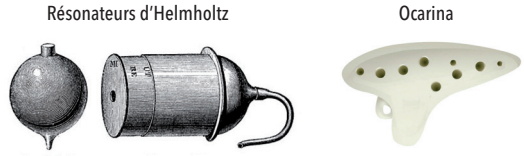
Source : Godwin J., 1980, *Athanasius Kircher, un homme de la Renaissance à la quête du savoir perdu*, Thames & Hudson, p. 71

De tels conduits ont été utilisés anciennement pour écouter indiscrètement ce qui se passait à distance comme le montre le dessin d'Anathasius Kircher. À petite échelle, le tube souple du stéthoscope remplit également cette fonction.

TERMES	DÉFINITIONS ET EXEMPLES
Ondes diverses (transversale, longitudinale) Figure G.15 Voir les sites ***sWR*** et ***sJW***.	 **Corde vibrante** Une corde excitée à l'archet génère au moins trois types d'ondes. Une onde *transversale* dont l'amplitude de déformation est perpendiculaire au sens de propagation le long de la corde ; c'est l'onde qui produit la « note » jouée (A). Une onde *longitudinale* ou onde de compression, qui progresse dans le sens de la corde, à une vitesse plus grande que l'onde transversale (B). L'excitation ne se produisant pas strictement dans un plan, il existe aussi des ondes de *torsion* (non représentées). Les fréquences et les amplitudes de ces trois types d'ondes contribuent au spectre du son rayonné par l'instrument. La figure de droite représente l'enveloppe de la déformation de la corde frottée pendant un cycle (onde transversale). **Tuyau sonore (C)** La propagation des ondes aériennes responsables du mode de vibration qui produit la hauteur fondamentale se fait dans l'axe du tuyau, c'est-à-dire dans le même sens que les oscillations locales des particules d'air. Ce sont des ondes *longitudinales*.
Oscillation	Terme employé pour désigner un mouvement vibratoire. En acoustique, on distingue les oscillations *libres* et les oscillations *entretenues*.
Oscillation libre ou entretenue	Un corps excité par une impulsion isolée vibre en émettant les fréquences de ses modes propres : c'est l'*oscillation libre*. La durée globale de la vibration dépend des caractéristiques d'amortissement du corps. Le son perçu pendant l'extinction varie au fur et à mesure de l'extinction des différents modes. C'est le cas de tous les instruments frappés comme la cloche, le steel-drum, la corde de piano. Lorsque la vibration est entretenue par le souffle ou par un archet, il s'établit un régime d'*oscillations entretenues* à la fréquence d'un mode propre particulier. La structure n'émet que les modes du système dont les fréquences peuvent se synchroniser sur la fréquence d'entretien du mode principal. Dans ce cas seulement, toutes les composantes sont harmoniques[4]. La durée du son cesse avec l'entretien. Tous les instruments à vent et à cordes frottées sont à placer dans cette catégorie. Voir chapitre 1, § 2.2 ; chapitre 2, § 4.3 ; **CK** p. 52.
Partiel (fréquence)	Terme général désignant une fréquence quelconque, un son partiel, dans le sens de partie d'un son complexe. Tous les sons non entretenus (cloche, piano, marimba) sont composés de partiels non harmoniques.
Partiel (mode vibratoire)	Fréquence d'un mode propre. On l'emploie aussi (par raccourci) pour désigner le son que produit un instrument vibrant dans un mode autre que le mode fondamental. Exemple : le partiel 2 d'une flûte traversière est à l'octave du mode 1 ou fondamental, alors que le partiel 2 d'une clarinette est à la dix-septième (un octave + une quinte) du partiel 1 ou fondamental. Voir chapitre 2, § 4.4, et § 4.5. Les harmoniques effleurés du violon ou de la guitare sont des partiels de la corde.
Partiel entretenu	Lorsqu'on entretient un partiel quelconque d'un tuyau ou d'une corde, le son produit est périodique : il possède une série harmonique.

TERMES	DÉFINITIONS ET EXEMPLES
Partiels successifs	La succession des *partiels* d'un tuyau conique à embouchure (comme le cor) est très voisine de la série harmonique. En musique, on emploie parfois le terme *harmonique* pour désigner ces partiels. En comparant les intervalles joués avec ceux de la série harmonique, on peut savoir quels sont les partiels produits par l'instrument. Si par exemple les deux premiers sons que l'on produit sur une trompette sont : *sib* et *fa*. Il s'agit des partiels 2 et 3 du tube, par analogie avec la quinte qui sépare les harmoniques 2 et 3. Dans beaucoup d'instruments à embouchure, le *fondamental* ou partiel 1, qui est le son le plus grave que peut produire le tuyau, sort difficilement et n'est pas utilisé en musique.
Partiels (justesse)	Par convention, on compare les intervalles entre partiels successifs à ceux des sons d'une série harmonique. Les partiels sont « justes » lorsqu'ils s'en approchent au mieux. Exemple : les partiels d'une flûte Boehm, d'un trombone ; les harmoniques effleurés d'une corde de violoncelle.
Période	La période T, ou durée d'un cycle complet, est l'inverse de la fréquence *f*. On a : T = 1/*f*. L'unité de temps utile change selon l'échelle des phénomènes (de la microseconde à l'année). Penser à la période du jour solaire, à celle des battements de cœur. Le si_4 dont la fréquence est voisine de 1000 Hz a pour période 1 milliseconde.
Périodique	Un phénomène est dit périodique lorsqu'il se reproduit identiquement à lui-même à intervalles de temps égaux. Il est caractérisé par sa *période*, son *amplitude* maximale et par la *forme* de la variation d'amplitude pendant un cycle. Dans le monde sonore, on considère deux ordres de phénomènes périodiques (ou quasi périodiques). 1/ les *cycles lents* : le tempo, les rythmes réguliers comme le tempo de la marche, certaines modulations. comme la modulation d'amplitude du vibraphone (4/s). 2/ les *cycles courts* : c'est-à-dire les sons proprement dits, dont la fréquence va de 20 à 20000 vibrations par seconde. Voir chapitre 3, § 3.
Périodique (mouvement)	Mouvement oscillatoire se reproduisant, identique à lui-même, à intervalles de temps égaux. En toute rigueur, un mouvement périodique n'a ni début, ni fin...
Périodique (son)	Son produit par une vibration entretenue (vibration forcée). Un son périodique est harmonique (par définition).
Phase Figure G.16 Voir les sites ***sJW ; sDR ; sWR***.	 Deux sons dont les vibrations sont synchrones sont en phase. Lorsqu'on effectue l'addition des amplitudes de deux signaux, il importe de tenir compte de leurs phases relatives. L'exemple de la figure ci-dessus, qui ne porte que sur deux harmoniques, permet de comprendre qu'il est parfois difficile d'interpréter la forme de l'onde résultante puisque celle-ci change d'aspect selon la phase des composantes, ce que ne perçoit pas l'oreille. Voir les entrées *Onde*, *Propagation*.

TERMES	DÉFINITIONS ET EXEMPLES
Propagation (réflexion, absorption, diffusion) Voir le site ***sJW***.	Considérons le déplacement d'une perturbation (onde) dans un milieu. À l'air libre, le son se propage par *ondes sphériques* ; dans un tuyau, le son se propage par *ondes planes*. En présence d'obstacles, une onde subit diverses transformations qui dépendent de l'état de surface de l'obstacle et du rapport entre les dimensions de l'obstacle et les longueurs d'onde des composantes de l'onde. Dans la *réflexion* (plan lisse et rigide), l'onde est renvoyée globalement dans une autre direction ; dans *l'absorption* (obstacle mou, souple), l'onde peut être intégralement « capturée » par le matériau des parois (chambre anéchoïque), aucun retour ne se produit ; dans la *diffusion* (murs sculptés, avec reliefs de dimensions différentes), l'onde est en partie réfléchie, ses composantes prenant des directions divergentes selon leur longueur d'onde. Voir **EL** chapitre 4 ; **TS** chapitre 3.
Rayonnement	Le rayonnement aérien d'un corps mis en vibration est une partie complexe de l'acoustique. Les ondes aériennes générées par un instrument à vent naissent aux orifices du tube (trous latéraux, pavillon). Dans les instruments à cordes, l'air est mis en vibration par les déplacements de la table d'harmonie et des parois de la caisse. S'y ajoutent les ondes se produisant dans la cavité de la caisse. La théorie simplifiée distingue les monopôle, dipôle et quadripôle. Voir **CK**, chapitres 12 à 14.
Rayonnement (monopôle, dipôle) Figure G.17 Voir les sites ***sDR***, ***sPF***, ***sJW***.	 Une source localisée peut être assimilée à un *monopôle* lorsque ses dimensions sont petites devant la longueur d'onde. Prenons l'exemple du volume d'air contenu dans la caisse d'une guitare. La première fréquence de résonance est de 120 Hz[5] ; la longueur d'onde λ = 340/120 = 2,8 m ; le diamètre du trou est de 0,80 m, ce qui est très petit par rapport à la longueur d'onde. Lorsque l'instrument vibre, le rayonnement du mode aérien de la cavité de la caisse est monopolaire. Une source de type *dipôle* comme la membrane d'un tambourin génère toujours deux ondes, l'une vers l'avant et l'autre vers l'arrière, qui sont en opposition de phase. Leurs amplitudes s'annulent lorsqu'elles se rencontrent sur les bords. On peut faire l'expérience suivante avec une cymbale suspendue (ou à défaut un couvercle de casserole). Approcher l'oreille très près (2 à 3 cm) du pourtour et déplacer la tête de part et d'autre du plan moyen de l'instrument pendant l'extinction du son. Au moment où l'oreille passe au niveau du bord, les composantes graves disparaissent, car elles sont en opposition de phase de part et d'autre de la surface ; elles réapparaissent lorsque l'oreille passe au-dessous du bord. Pour accroître le rayonnement aux basses fréquences d'un dipôle, il faut donc isoler l'avant de l'arrière. La pratique du *baffle* déclinée de diverses façons dans la construction des enceintes acoustiques se retrouve dans les instruments de musique. Les tambourins on généralement un rebord. Voir **CK**, p. 229.
Rayonnement (quadripôle) Figure G.18 Voir les sites ***sDR*** et ***sJW***.	 *Quadripôle.* Excité sur son premier mode vibratoire, un verre se déforme alternativement selon deux diamètres orthogonaux et rayonne dans les directions correspondant aux zones ventrales indiquées par « V » sur la figure G.18. Les zones marquées « N » n'émettent pas de son. Lors d'une excitation au doigt, le système modal se déplace en tournant, si bien que l'amplitude change dans le temps pour un observateur fixe (et en particulier pour le microphone). Les variations de l'amplitude sont très audibles et clairement visibles sur l'analyse de l'amplitude. Voir chapitre 1, figure 1.3 ; chapitre 2, figure 2.15. Voir aussi **CK** (chapitre 12, § 4.5).

TERMES	DÉFINITIONS ET EXEMPLES
Réflexion	Le mouvement résultant de la propagation d'une onde dans un système vibrant (tuyau, corde) dépend des conditions aux extrémités, notamment selon le fait que celles-ci sont fixes ou libres. Voir **EL** chapitre 15. Sites : ***sJW*** ; ***sWR***.
Registre	Terme musical désignant une succession de sons de la tessiture d'un instrument produits par un système vibratoire donné et qui présentent une bonne homogénéité de sonorité. L'exemple le plus clair est celui de l'orgue. On connaît aussi les registres de la clarinette : du grave à l'aigu, on différencie les sons du chalumeau, du medium et du clairon (termes qui changent selon les ouvrages). Les chanteurs distinguent différents registres dans leur voix. Le nombre et les dénominations des registres des chanteurs sont très variables d'un individu à l'autre (voir Roubeau & coll., 2009).
Résonance	L'excitation par résonance est une excitation indirecte qui se produit par voie aérienne ou par couplage solide. Soit un corps (A) vibrant à la fréquence *f* et un autre corps (B) possédant des modes propres de fréquences variées. Pour qu'il y ait résonance, deux conditions sont nécessaires. 1 : la fréquence d'excitation de A doit être très voisine, voire identique à celle d'un des modes propres de B ; 2 : l'apport d'énergie doit être en phase avec l'oscillation de B. Soumis à une telle excitation, le système va être le siège d'oscillations de plus en plus importantes, jusqu'à atteindre un régime d'équilibre qui dépend des éléments dissipatifs propres au système, ou de la rupture d'un de ses composants. La résonance démarre toujours avec une certaine latence, ce qui prolonge la durée du transitoire. Employé pour qualifier une salle d'écoute, le mot résonance prend un tout autre sens. Il signifie que le son continue à se propager après qu'a cessé l'excitation. Dans ce cas, le terme correct est plutôt *réverbération*.
Résonateur Figure G.19 Koenig, R., 1889, figures 15 et 16, [Virtual Laboratory] © ocarina : Artstada/ Shutterstock	 Ce terme a au moins deux sens dans la littérature acoustique.. 1/ Dans la description fonctionnelle d'un instrument de musique, on sépare l'excitateur et le résonateur. Ici le terme résonateur désigne la partie de l'instrument dont les modes propres d'oscillation vont être sollicités par l'excitateur (archet, anche, etc.). 2/ Du point de vue physique, un résonateur est un corps creux quasi sphérique communiquant avec l'air extérieur par une ouverture de petites dimensions. L'exemple le plus connu est celui du résonateur dit de Helmholtz. On désigne par « fonctionnement résonateur » tout mode vibratoire d'une cavité dont la longueur d'onde est très supérieure à sa dimension interne la plus grande. C'est le cas de la bouteille et, en musique, de l'ocarina. Les paramètres déterminant la fréquence principale d'un résonateur sont : le volume intérieur de la cavité (V), la section de l'orifice (s) et la célérité de l'air (c). Interviennent aussi la géométrie du raccordement du *col* à la sphère. Un résonateur ne produit pas d'énergie : il la répartit différemment dans le spectre. Voir le § 4 de l'annexe C, pour un exemple numérique de calcul appliqué à une bouteille. Voir aussi le chapitre 9, § 2 pour les résonateurs de la voix chantée.
Son	Dans cet ouvrage, le terme *son* désigne tout phénomène vibratoire perceptible par les oreilles (son musical, bruit, ou autre), qu'il soit transmis au système auditif par l'air, par l'eau ou directement par les os du crâne.
Son pur, son complexe	Le son *pur*, simple ou sinusoïdal, est une fréquence isolée. C'est un signal de laboratoire utilisé en physique et en psychophysique. Parmi les sons réels, le sifflet oral qui excite la résonance principale de la cavité buccale peut être considéré comme un son pur, donc simple. Un son *complexe*, composé de plusieurs fréquences pures, peut être harmonique ou inharmonique. En psychoacoustique, l'opposition simple-complexe, importante pour caractériser les traitements des signaux sensoriels, ne recouvre pas nécessairement le sens physique. Par exemple, la sensation de hauteur d'un son complexe riche en harmoniques, qui est plus immédiate et précise aux oreilles d'un auditeur, paraîtra plus « simple » que celle d'un son pur.

TERMES	DÉFINITIONS ET EXEMPLES
Spectre	Représentation de l'amplitude du son en fonction de la fréquence. Les différents types de spectres : le spectre *instantané*, le spectre *moyenné* sur une longue durée (LTAS) rendent compte différemment des caractéristiques des sons. Voir chapitre 2, § 3.5. En toute rigueur, le spectre s'exprime dans l'espace des complexes. On ne s'intéresse ici qu'au module (l'autre partie étant la phase).
Source	Au sens de source sonore, ce terme peut désigner au moins deux choses : – la structure (ou générateur) à l'origine de la production du son (corde vibrante, jet oscillant, anche, cordes vocales), donc de l'oscillation ; – l'origine du rayonnement (trous latéraux d'une flûte, membrane de la timbale, orifice buccal). Voir **CK**, p. 688, note 13.
Top-down	En français, « de haut en bas ». Expression employée pour signifier l'influence des contrôles effectués par un auditeur au cours de la réception d'un son : contrôles s'effectuant depuis les centres supérieurs vers les organes sensoriels et les traitements neurologiques intermédiaires, via des circuits *efférents* (voir chapitre 3, figure 3.1).
Transitoire	Terme désignant des variations rapides, comme celles du début et de l'arrêt de la production des sons. Intervalles de temps très courts pendant lesquels les variations de la vibration sont très complexes, et nécessitent des techniques spécifiques d'analyse en traitement du signal.
Unités	Principales unités utilisées en acoustique. Le **Hz** (Heinrich Hertz) unité de fréquence, est le nombre de périodes dans une seconde. Le **Bel** ou bel, (Graham Bell), unité de mesure de l'intensité, est le logarithme (à base 10) du rapport entre deux valeurs de la puissance ou de l'amplitude. En pratique, on utilise le **dB** (décibel), 10 fois plus faible. L'ajout d'un suffixe signale un dB défini pour un usage précis : dB A, dB SPL. Le **cent** et le **savart** sont des unités de mesure logarithmique des intervalles musicaux. Voir chapitre 8, page 406.
Vibration	Mouvement de va-et-vient autour d'une position d'équilibre. Voir aussi *Oscillation*.
Vibration forcée	Vibration entretenue par un apport continu d'énergie : souffle, frottement. L'excitateur et la structure sont couplés de sorte que la vibration produite se synchronise sur la fréquence d'un mode donné de la structure et fournit l'énergie en bonne concordance de phase avec celle-ci.
Vibration libre	Vibration produite par une impulsion isolée, un apport ponctuel d'énergie (choc). La structure vibre sur ses modes propres (modes libres).
Zone d'écoute dominante	Zone des fréquences situées entre 700 et 1 800 Hz qui joue un rôle particulièrement important dans la perception de hauteur des sons inharmoniques (cloches) et dans le phénomène dit de *fondamental virtuel*. Voir chapitres 3 § 2.9 et 6 § 3.5.1. Les limites de cette zone d'écoute cognitive varient légèrement selon les catégories de sons concernés et selon le contexte sonore.

1. La célérité du son (c) = 20 $(T + t)^{1/2}$ avec (c) en mètres, (T) = Température absolue (−273 °C) et (t) = température du gaz.
2. ko pour kilo octet (bytes pour les Anglo-Saxons). 1 octet = 8 bits ; 1 kilo octet = 8×2^{10}, soit 8×1024 octets. La seconde de son stéréo occupe $(44100 \times 2 \times 16)/(8 \times 1024)$ = 172,26 ko de mémoire.
3. La durée d'un échantillon est l'inverse de la fréquence d'échantillonnage. Pour *fe* = 44,1 kHz, la durée est de 22,67 $\times 10^{-6}$.
4. Dans la réalité instrumentale, on rencontre plusieurs bruits se superposant au spectre périodique : bandes de souffle, fréquences inharmoniques au transitoire d'attaque.
5. Mesure effectuée sur une guitare Suzuki.

ANNEXE H

Contenu du DVD-Rom d'accompagnement

Comme indiqué dans l'avant-propos de l'ouvrage, nous vous proposons sur le DVD-Rom d'accompagnement les exemples sonores de l'ouvrage sous deux formats : des « livrets-sons » au format ePub 3 pour écouter les sons sur son Smartphone ou sa tablette en parallèle à la lecture du livre papier et le livre complet au format PDF avec sons intégrés, à lire et écouter sur sa tablette ou sur son ordinateur.

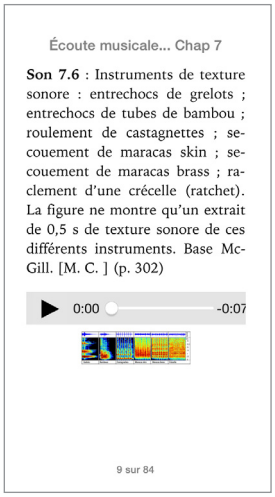
Ces choix méritent quelques explications, car vous vous demandez sans doute pourquoi nous n'avons pas tout simplement fourni directement les sons sous forme de fichiers audio au format MP3 ou WAV. Nous avons constaté que la plupart des lecteurs MP3 et des applications de lecture audio sur Smartphone enchaînent automatiquement les morceaux, sans qu'il soit possible de paramétrer un arrêt à la fin de chaque séquence sonore. Les sons de l'ouvrage ayant une durée très courte, de quelques secondes à quelques dizaines de secondes, il serait extrêmement laborieux d'avoir à mettre son appareil en pause après l'écoute de chaque son pour poursuivre la lecture du livre jusqu'au son suivant. Une playlist rassemblant tous les sons de l'ouvrage se révèle ainsi impraticable.

1. Livrets-sons au format ePub

Nous avons donc choisi d'intégrer les sons dans des fichiers eBook au format ePub 3 que nous avons baptisés « livrets-sons » (un par chapitre). Ces derniers peuvent être lus sur Smartphone ou tablette à l'aide des applications recommandées dans le tableau ci-dessous.

APPAREIL	APPLICATION DE LECTURE RECOMMANDÉE	CONFIGURATION MINIMALE
iPod/iPhone/iPad	iBooks (gratuite)	iPod 4 ou version ultérieure, iPhone 3G ou version ultérieure, iPad 2 ou version ultérieure, iOS 6 ou version ultérieure
Smartphone ou tablette Android	Gitden Reader (gratuit) Moon+ Reader (gratuit)	Android 4.0 ou version ultérieure

Nous proposons, en plus des fichiers ePub découpés par chapitre et donc de poids relativement légers, un fichier ePub rassemblant les sons de l'ensemble des chapitres. Nous vous recommandons de n'utiliser ce fichier global que sur un appareil récent, par exemple un iPhone 5 ou 6, car le poids du fichier est tel qu'il met un temps prohibitif pour s'ouvrir dans l'application iBooks sur des iPhone plus anciens. Il en est de même en environnement Android.



La capture d'écran ci-contre montre comment se présente l'interface de ces fichiers ePub : une page par son, avec un descriptif du son qui reprend la légende du son située en fin de chapitre, une barre de lecture audio sur laquelle il vous suffit de cliquer pour déclencher la séquence sonore, et une vignette du sonagramme ou de l'illustration associée au son, qui sert de repère visuel. Le format audio utilisé pour les sons est le format MP3.

Pour une expérience de lecture/écoute optimale, nous vous conseillons de vous installer confortablement avec le livre papier et votre Smartphone équipé d'un casque ou, encore mieux, connecté à votre chaîne hi-fi.

2. Livre complet au format PDF avec sons

Si vous êtes adepte de la lecture sur écran, nous vous proposons encore mieux : une édition numérique du livre complet au format PDF avec sons intégrés.

Pour la lecture sur ordinateur, vous devez ouvrir le fichier avec Adobe Acrobat Reader (gratuit), seul logiciel reconnaissant les fichiers sons. À la première lecture d'un son, un avertissement s'affiche. Ouvrez « Options » et cliquez sur « toujours approuver ce document ».

Pour la lecture sur tablette, vous pouvez utiliser l'une des applications recommandées dans le tableau ci-dessous.

APPAREIL	APPLICATION DE LECTURE RECOMMANDÉE	CONFIGURATION MINIMALE
PC	Adobe Acrobat Reader (gratuit)	Windows XP ou version ultérieure, Adobe Acrobat Reader 6 ou version ultérieure
Mac	Adobe Acrobat Reader (gratuit)	Mac OS X 10.0 ou version ultérieure, Adobe Acrobat Reader 6 ou version ultérieure
iPad	PDF Expert de Readdle (9,99 € pour la version 5) ezPDF Reader de Unidocs (3,99 €)	iPad 2 ou version ultérieure, iOS 6 ou version ultérieure pour ezPDF Reader et PDF Expert 4, iOS 7 ou version ultérieure pour PDF Expert 5
Tablette Android	ezPDF Reader Pro de Unidocs (3,22 €)	Android 4.0 ou version ultérieure
Note. Les versions iOS et Android de l'application Acrobat Reader d'Adobe, l'éditeur de logiciels à l'origine du format PDF, ne supportent malheureusement pas les PDF avec sons, ce qui nous a obligés à proposer des solutions alternatives payantes pour ces deux plateformes.		

Deux versions vous sont proposées : l'une avec sons au format MP3, l'autre avec sons au format WAV pour une meilleure qualité d'écoute. En raison de son poids, nous avons dû découper le fichier PDF avec sons au format WAV en trois fichiers : le premier inclut les sons de l'introduction au chapitre 3, le deuxième ceux des chapitres 4 à 6, et le dernier les sons des chapitres 7 à 9.

3. Compléments vidéo

Les analyses sonagraphiques reproduites dans le livre sont des images statiques. Les vidéos que nous vous offrons permettent d'associer en temps réel la forme visuelle des sons et leur écoute, en suivant le curseur. Avec un peu d'habitude, vous pourrez même anticiper l'écoute à la vue des images.

Les sonagrammes ont été réalisés avec Audiosculpt v.3.3.8 et leur lecture sonore captée avec QuickTime.

1 – Séquence urbaine (durée : 20'') – Introduction, page 3.

2 – Figures de Chladni (durée : 2'05) – Chapitre 1, page 15.

La vidéo « Chladni » visualise les vibrations d'une plaque rectangulaire excitée par un son sinusoïdal continûment ascendant, émis par un haut-parleur situé sous la plaque. Le sel se rassemble le long des lignes nodales qui sont immobiles.

3 – Analyse d'une cloche (durée : 2') – Chapitre 1, page 24 – Son 1.14, cloche lorraine.

4 – Musique improvisée au *pakhavaj* (durée : 60'') – Chapitre 4, page 197, figure 4.45. Voir légende du son 4.40, page 205.

5 – Musique pour pianoforte, *Polonaise en sol mineur* de F. Chopin (extrait) (durée 1'17) – Chapitre 4, page 197, figure 4.45. Voir légende du son 4.41, page 205.

6 – Musique composée à l'ordinateur - Extrait de *Mutations*, de J.-C. Risset (durée : 1'07) – Chapitre 4, page 197, figure 4.45. Voir légende du son 4.42, page 205.

7 – Guimbarde et chant sigit (durée 57'') – Chapitre 6, page 248, figure 6.19. Voir légende du son 6.25, page 283.

8 – Arc ngbaka et chant (durée 1'13) – Musicien : N. Masémokobo, République centrafricaine – Chapitre 8, page 402, figure 8.12. Voir légende du son 8.6, page 439.

9 – Paysage sonore (durée 1'35) – Scène enregistrée à l'abbaye d'Ardenne, Normandie, juin 2010. *Une personne descend l'escalier, passe deux portes, marche sur le gravier à l'extérieur et se dirige vers le verger ; on entend des musiciens jouant dans une salle fermée puis les oiseaux qui sont de plus en plus nombreux (rouge-gorge, fauvette, pigeon ramier). La circulation automobile (route à 400 m) est continue.*

4. Structure du DVD-Rom

Le schéma ci-dessous détaille l'arborescence des répertoires du DVD-Rom. Le poids de chaque fichier est indiqué entre parenthèses.

Lisez-moi.pdf

Livrets-sons au format ePub 3

Livrets-sons par chapitre

Livret-sons_Intro.epub (2,2 Mo)
Livret-sons_Chap1.epub (33 Mo)
Livret-sons_Chap2.epub (19 Mo)
Livret-sons_Chap3.epub (19 Mo)
Livret-sons_Chap4.epub (34 Mo)
Livret-sons_Chap5.epub (14 Mo)
Livret-sons_Chap6a.epub (30 Mo)
Livret-sons_Chap6b.epub (23 Mo)
Livret-sons_Chap7.epub (37 Mo)
Livret-sons_Chap8.epub (26 Mo)
Livret-sons_Chap9.epub (17 Mo)

Livret-sons de l'ouvrage complet

Livret-sons_OuvrageComplet.ePub (243 Mo)

Livre au format PDF avec sons

Avec sons MP3

LivrePDF_SonsMP3.pdf (386 Mo)

Avec sons WAV

LivrePDF_SonsWAV_Part1.pdf (425 Mo)
LivrePDF_SonsWAV_Part2.pdf (527 Mo)
LivrePDF_SonsWAV_Part3.pdf (457 Mo)

Compléments vidéo

Video1-Séquence urbaine.mov (8,4 Mo)
Video2-Chladni.mp4 (21,6 Mo)
Video3-Cloche.mov (71,8 Mo)
Video4-Pakhavaj.mov (50,8 Mo)
Video5-Pianoforte.mov (58,9 Mo)
Video6-Mutations.mov (47 Mo)
Video7-Guimbarde-Chant sigit.mov (35,3 Mo)
Video8-Arc musical.mov (27,9 Mo)
Video9-Paysage sonore.mov (38,6 Mo)

Bonne lecture et bonne écoute !

Index

A

accordage (polyphonie, sons fixes) 420–435
- bases acoustiques 421–422
- commas 422–425
- systèmes d'accordage 425–432

accordage pratique
- complexité des sons réels 432–433
- compromis harmonie/mélodie 435
- piano 433
- quinte pure au clavecin 438
- tempérament égal 431
- tempéraments inégaux 427

amplitude 19, 515
- enveloppe temporelle 47
- variations temporelles et rayonnement 48

analyse acoustique
- analyse-synthèse 52
- choix des paramètres (voix) 67
- comparaison de logiciels 71
- étapes 35
- historique 74
- interprétation des mesures 56, 68, 69
- représentations groupées 56, 58
- spectrotemporelle 90, 156
- temporalité 39
- terminologie 46
- tridimensionnelle 34

anamorphose 168, 180, 200, 324
anticipation 146–149, 198
arc musical 27, 402, 465
attention auditive
- alertée, mobilisée 144, 170, 176, 189, 277
- sélective 90, 113, 147, 149, 162, 226, 343, 353, 361
- *Voir aussi* anticipation

audiogrammes
- pertes avec l'âge 96
- tonal, verbal 96

B

bandes critiques 98
- analyse fréquentielle (Barks, ERBs) 100
- fréquence et intensité 99
- largeur selon la fréquence 99
- modulations 121
- perception des modulations 124

bandes ventriculaires. *Voir* larynx
Barks (bandes critiques) 99
battements 120
- binauraux 131
- effet esthétique 415
- harmonique commun 413
- perception et tessiture 124
- richesse spectrale (écoute) 413
- sons voisins purs, complexes 411
- sons voisins très lents 413

bottom-up 142, 151, 165, 516
brillance 209, 211, 242, 356, 357, 359, 373, 383

C

cabrette auvergnate et grelots 224
catégorisation prototypique 151, 182–193
- catégorisation libre 185, 191
- niveau de base 184
- prototype 184, 192
- similarité 141, 184, 186

cents, savarts
- calcul d'un intervalle 406
- tableau récapitulatif des intervalles 479

chant diphonique. *Voir* chant harmonique
chant harmonique 26, 245
- chant de F1 464
- chant de F1/chant de F2 462
- F1 et F2 en octave 465
- femme xhosa 248
- kargiraa 465
- mongol 398
- tuva 248, 465

chevrotement 122
chimère
- écoute 309
- schéma de reconnaissance 320

cloche de volée
- analyse des partiels (écoute) 24
- cloche à deux notes 263
- note au coup 257
- note de quarte 259

cloche tube 261, 279
cochlée
- cellules ciliées (internes et externes) 88
- membrane basilaire 87

codages de l'information sensorielle 91
- codage spatial ou tonotopique 91
- codage temporel 93

commas
- calcul en cents 481
- enharmonique 425
- pythagoricien 424
- syntonique 422

contour mélodique 234, 277
contrebasson 240
cor des Alpes 28, 44
cordes vocales. *Voir* larynx
crescendo musical 210
- clavecin 216
- piano 218
- saxophone 212
- trompette 211

D

diapason
- modes vibratoires 16

diapason d'accord
- variabilité historique 435

différentiels (sons) 123, 125, 415
- duo de flûtes 416

discrimination auditive (carte) 106
discrimination temporelle 125
- seuils d'intégration 126
- transitoires 126

dynamique d'un instrument de musique
- mesure en dB 214
- variations avec la tessiture 214
- variations par la registration 216

E

échelles de sélection d'harmoniques
- arc musical 402, 407
- chant harmonique 398
- guimbarde du Kazakhstan 400

écoutes
- acte d'écoute (Husserl) 146
- causale, événementielle 144, 146, 199
- de premier niveau (sources) 199
- de second niveau (séquences) 199
- musicales 195
- qualitative, choisie 143, 144, 146, 167, 199, 366
- tests d'écoutes 189

écoutes du timbre
- écoutes du son 368
- instrumentistes, luthiers, compositeurs 368
- qualités et identité 366

effet Haas 132
électroglottographie - EGG 449
émergence sonore
- contraste de forme 170

émergence spectrotemporelle
- cabrette 224
- métro 223
- vibrato 225

enveloppe d'amplitude 331
- battements 412, 516
- de la déformation 526
- de la forme d'onde 253
- des harmoniques 319

enveloppe dynamique 48, 252, 255, 297, 306, 332, 365
enveloppe spectrale 50, 234, 242, 279, 314, 322, 324, 326, 356, 357
- formants 315, 317, 518

enveloppe temporelle 41, 47, 123, 157, 444
excitation 7
- impulsion ou entretien 9

F

FFT
  Voir transformée de Fourier (FFT)
flatterzunge 123
flûte traversière 312, 315, 322, 327, 330, 337
  flûte de Quantz 428
formant du chanteur 209
formants 50, 67, 158, 299, 318
formants vocaliques 324, 325, 453
formant-tuning 458, 460, 461, 471
forme d'onde 41, 42
  sinusoïdale, carrée, en dents de scie 56
forme spectrotemporelle 152, 157, 179, 200
forme temporospectrale 157, 171, 291, 311, 343, 344, 385, 444, 445
formes et catégories 200
formes sonores 70, 154
  collection A 155, 156
  collection B 155, 156, 187
  de second niveau 160
  dimension temporelle 154, 158
  émergence forme-fond 171
  émergence sur le fond 170
  forme et matière 158, 159, 160, 162, 173
  musique 162, 171, 196
  niveau de structuration 161
  parole 161, 168, 179
  premier et second niveau 160
  reconnaissance 154
  signaux d'avertissement 159
  sources et séquences 159
  typologie 157
  variabilité 166–167
fréquence
  codage spatial 91
  codage temporel 93
fréquence d'échantillonnage 44, 56, 67, 72
fréquence de Helmholtz 307
fry. *Voir* mécanismes laryngés

G

galoubet et tambourin 225
gamme
  de Pythagore 421
  de Zarlin 405, 421
  naturelle 392, 405
Gestalttheorie 152
  Voir aussi théorie de la forme
glotte. *Voir* larynx
gong birman 48
guimbarde 400

H

harmonique 7 398–400, 403
harmonique commun 413
  schéma 54, 401
harmoniques
  intervalles de la série 22
  série 22
harmoniques (écoute)
  basson 22
  harmonica 22
  trombone 22
  verre percuté et frotté 21
harmoniques (série)
  théorie musicale 397
harmoniques effleurés 16, 26, 29, 395
harpe 11
hauteur des sons
  dimension cognitive (test voix-sifflet) 235
  hauteur et timbre 240, 251, 253, 264, 274, 278, 279
  instabilités des sons réels 265
  perception des sons vibrés 268
  qualités de sons connus 279
  sons périodiques 237–251
  structure acoustique 230–234
  vibratos, trilles 265–276
  zone mixte (graves) 118, 119
  Voir aussi hauteur virtuelle
  Voir aussi sensation de hauteur
hauteur des sons apériodiques 251–264
  cas du piano 252
  cloche à 2 notes (saillance temporelle) 263
  cloche tube 256
  diapason à fourche, marimba 254
  modèle harmonique 255
  note de la timbale 260
  notes d'une cloche d'église 257
  sons successifs (paradoxe) 262
  sons successifs très amortis 264
  sons successifs, cloche tube 261
  zone d'écoute dominante 261
hauteur des sons périodiques 237–251
  hauteur spectrale/tonale 240
  incidences du contenu spectral 239
  limites grave/aigu 237
  singularité des sons purs 237
hauteur spectrale/tonale
  bilan (schéma) 276
  conflits d'écoute 245, 248
  contrebasson (test) 240
  formants de la voix humaine 244
  sons paradoxaux 242
  voir chant harmonique (diphonique) 245
hauteur virtuelle 116
  note d'une cloche tube 256
  note de quarte d'une cloche 259
  quintina 467
HRTF 132

I

icophone 162, 168, 179
illusions auditives 141, 153, 180
impulsions
  analyse FFT 44
  impulsion ou entretien 9
  propagation dans un tuyau 13
impulsions successives
  discrimination temporelle (TFT) 125
  hauteur perçue 118, 119
intégration temporelle
  impulsions successives 119
  modulations 121
    intensité des sons musicaux 207–227
  chanteur 209
  contexte temporel 220
  crescendo 211
  émergence spectrotemporelle 223, 224
  enrichissement spectral 211
  saxophone 212
  sources connues, catégorisées 220
  sourdines 213
  téléphone 209
  zone des 3 kHz 208, 209, 210
  Voir aussi nuances musicales
  Voir aussi sensation d'intensité
intervalles harmoniques
  voix chantée 403
intervalles mélodiques
  arc musical 402, 407
  flûte traversière (Haydn) 408
  hautbois (Berlioz) 409
intervalles musicaux
  agrandis en échelle linéaire 55
  cadre intervallique 394
  catégories et tolérance 391
  coïncidences d'harmoniques 54
  données perceptives 391
  interprétation des mesures 404, 410
  justesse et variabilité 408
  monodie, polyphonie 393, 411
  octave, quinte et quarte 394
  perception 408
  perception d'intervalles étrangers 398
  traditions orales 395
  Voir aussi tableau des intervalles 479
intervalles, mesures
  calcul en cents et en savarts 406, 479
  discussion 404
  rapports numériques 396, 405
  recherche de l'harmonique commun 401
  tableau récapitulatif 479
IRM, fIRM (imagerie) 90, 452

L

larynx 446
  cartilages 447
latéralisation auditive 128
  seuil temporel binaural 131
localisation auditive 128
  effet de précédence 132
  en situation réelle 132
  plan horizontal 129
  plan horizontal (discrimination angulaire) 130
  plan vertical 131
localisation des sons dans l'espace 127

M

masquage 92, 101–102
  simultané 99
  test d'écoute 102
mécanismes laryngés 450
mel 110
mélodies d'harmoniques 26, 245
  arc musical 27, 247
  chant 26, 66, 248
  guimbarde 27, 247
mélodies de partiels 28
  cor des Alpes 28
  flûte tilinca 28, 65

monocorde vietnamien 29
trompe de chasse 28
trompette marine 29
tuba 66
membrane basilaire
codage spatial (tonotopique) 87
mesures de la fréquence 40
échelle logarithmique, linéaire 53
estimation manuelle 41
méthode temporelle 42
par autocorrélation 42
partiels d'un son inharmonique 59
son grave 44, 45
mode propre
déformée modale 14
fréquence propre 14
géométrie du corps vibrant 14
modes vibratoires 522
diapason 16
partiels 526
visualisation de Chladni 15
modulations (amplitude, fréquence) 120–125
effets perceptifs de la rapidité 122
lentes (battement, vibrato) 121
rapides (domaine des hauteurs) 123
monocorde 395
monocorde vietnamien 29
mots du timbre
brillance 357, 359, 373, 383
brillance (tests) 356
description vocalique 373
Helmholtz 353–355
rugosité 340, 356, 383
son timbré (flûte, clarinette, voix) 371
traités de musique 351
Voir aussi parler du son
multiphoniques
flûte traversière, hautbois 418
PGCD 419
zone d'écoute dominante 420

N

non-linéarités 94, 211
nuances musicales
long terme et court terme 221
modifications spectrotemporelles 217, 218
notation sur partitions 221
silences d'articulation 220
sources connues, catégorisées 218
vibrato 217

O

octaviation, quintoiement 25, 63
octobasse 214
oiseaux 68, 69
onde
longueur d'onde 18, 522
notion 12, 523
propagation 12, 18, 524
stationnaire 14
transversale, longitudinale 526
oreille
aire audible 95
analyse de la fréquence 98
masquage fréquentiel 101
seuils absolus 95
zone sensible de l'oreille 106
oreille absolue 437
oreille externe 82
directivité 82
effet coupe-vent 83
oreille interne
canal cochléaire 87
cochlée ou limaçon 85
fenêtre ovale, fenêtre ronde 86
membrane basilaire 87
rampes vestibulaire et tympanique 86
système vestibulaire (équilibration) 85
oreille moyenne 83
chaîne des osselets 83
fonctionnalités 83
trompe d'Eustache 85
organisation perceptive 143, 173–178
grandeur des intervalles 174
polyphonie (bonne continuation) 176
polyphonie (localisation spatiale) 178
rôle du timbre 176
sons successifs 173
orgue
jeu de tierce (timbre) 322
ornements mélodiques
flûte traversière 274
oscillations 526
amplitude 12
fréquence 217, 269, 271, 312, 314, 331
vibrations 19, 20, 253, 529
oscillations entretenues 526
oscillations libres 19, 25, 26, 74, 515, 526

P

parler du son
catégorisation libre et verbalisations 376
vous avez dit clair ? 378
partiels 526
justesse 527
partiels et harmoniques 25
suite des partiels (tuba) 25
partiels d'un son percuté
analyse d'une cloche 24
analyse du steel-drum 58–59
analyse du verre 62
partiels successifs (modes propres)
flûte traversière, clarinette 63, 64
tuba 66
pattern d'excitation 101, 102
perception 140–143
anticipation, prévisibilité 142, 146, 148, 165, 189, 198
catégorielle 176, 193
écologique 141, 143, 149, 198
en contexte, située 148, 198
identification ou qualification 144
organisation perceptive 143
physicalistes et cognitivistes 142
schéma fonctionnel 147, 165, 198
sélective 142, 147
temporalité 144, 152
perception catégorielle (test d'écoute) 391
percussions
scène sonore 70
typologie *Voir* timbre 300
period-doubling 26, 447, 464, 465
période 19, 41, 59, 72
phones 108
propagation 528

Q

quintina (chant sarde) 467
expérience de filtrage (écoute) 468

R

rayonnement 8, 17, 528
résidu 116
résolus (sons) 100, 114, 115
résonateur de Helmholtz 529
roulement 123
rugosité 301, 340, 356, 383

S

scène sonore, paysage sonore
conversations, rires, musique 170
dans le métro parisien 223
nuit en Vendée 69
percussions 70
scène urbaine 3
sensation
loi de Fechner-Weber 103
réponse cellule sensorielle 92
transduction nerveuse 91
sensation (et logarithme) 103
sensation d'intensité (sons complexes) 113
sensation d'intensité (sons purs) 108
isophonie 108
isosonie 109
rôle de la durée 109
sensation de hauteur 10, 35, 38, 41, 59, 87, 98, 406, 410, 436, 469, 529
sensation de hauteur (sons complexes) 113
hauteur tonale, hauteur spectrale 117
hauteur virtuelle 116
limite inférieure 119
théories temporelles 116
théories tonotopiques 115
sensation de hauteur (sons purs) 110
intervalles 111
limite aiguë 112
octaves et tessiture (test) 111
rôle de l'intensité 110
rôle de la durée (test) 110
seuils absolus
seuil d'audibilité, seuil de douleur 95
seuils différentiels 101–106
d'intensité, test d'écoute (800 Hz) 106
de fréquence, test d'écoute (1000 Hz) 106
signaux d'avertissement
métro 223
sifflet 208, 224
similarité 141
situation d'écoute
conversation 5

environnement 5
musique 6
son
apériodique, inharmonique 24
complexe 24
complexe (psychoacoustique) 113
entretenu, impulsionnel 9, 11
fondamental 24
partiel 24
périodique, harmonique 24
production 7
pur, sinusoïdal 24
sonagramme 2, 37
échelle linéaire, logarithmique 53
fenêtre d'analyse 44, 72, 119
lecture 38
sones 109
sonification 179
sources instrumentales (timbre)
sources de hauteur définie 308–322
sources impulsionnelles 300–308
typologie spectrotemporelle 297–314
spectres
à court terme 49
enveloppe spectrale 50
moyennés IDS (orgue) 51
moyennés LTAS (hautbois, flûte) 50
spectrotemporelles 200
steel-drum 58
streaming. *Voir* organisation perceptive 173
système auditif 79
hémisphères cérébraux 90
périphérique 81
voies nerveuses 89
systèmes d'intervalles
cadre intervallique (arc musical) 407

T

tampoura (jouari) 225
téléphone 210
tempérament égal 411
piano 434
référence 406
tableau des fréquences par cent 483
théorie de la forme 151, 152–170
forme et fond 169–173
forme et matière 152, 153, 159–165, 173, 179
formes fortes, faibles 153
organisation perceptive 143, 173–178
variabilité des formes, anamorphoses 166–168
tierce harmonique (6/5) 403
tilinca roumaine 28, 65
timbale 260
timbre 233, 234
attributs perceptifs 356
définitions acoustiques 296
du timbre au son 296, 368
écoutes identitaire et qualitative 290
étude du timbre (deux approches) 384
historique 289
timbré 353, 371
timbre et musique 291–296
timbre identitaire 297–345
classification automatique 344
cohérence en musique 338
écoute mélodique et timbrale 339
forme temporospectrale 343
oreille absolue 437
permanence-variation 332
tests de reconnaissance 309
timbre de l'instrument 332
timbre de la note 307, 311
timbre et tessiture 322
typologie acoustique 297–311
typologie instrumentale 342
typologie perceptive 308
variabilité 322–336
timbre qualitatif
catégorisation identitaire (timbré) 371
prototype de guitare baroque 349
prototype de référence 379
Voir aussi mots du timbre 351
timbre, catégorisation 346–350
familles d'instruments (flûtes, cordes) 347
niveau de base, expertise 346
prototype de référence 349
timbre, instrument
classification acoustique 342
gammes chromatiques 326
nouveaux modes de jeu (flûte) 337
souffle, bruits 336
variabilité de l'interprète 333–338
variabilité en tessiture 322
timbre, sources entretenues 318
caractéristiques spectrales 314–319
formant 315, 317, 319
sourdines de trompette 319
tests de reconnaissance 309
timbre de la note 307, 311
transitoires d'attaque 311, 331, 333, 334
typologie acoustique 308
typologie perceptive (test) 310
voix humaine 318, 341
timbre, sources impulsionnelles
cordes pincées, frappées 304
hauteur déterminée 302
hauteur indéterminée 300
tube, lames, membranes 303
timbre, tests d'écoute 356–364
catégorisation libre 376
évaluation d'instruments de musique 379
méthode d'étude 375
sons de synthèse (200 Hz) 357
test multi-instruments Grey (mib_3) 359
tests mono-instrument (Gaillard, Bensa) 363
validité écologique 375
voix chantée 380
top-down 142, 151, 165, 375, 530
transformée de Fourier (FFT) 40, 519
dualité temps-fréquence 43, 44, 119
fenêtre d'analyse 45
paramètres de l'analyse (exemples) 72
tableau des relations ($\Delta f.\Delta t$) 72
transitoire d'attaque 321, 343
analyse dynamique 47
analyse temporospectrale 311
articulation 334
durée 313, 321
rôle de l'interprète 333
steel-drum 58
tessiture 331
tuyau d'orgue (écoute) 52
trémolo 122, 123
trille instrumental ou vocal 272
trille vocal 273
trille et vibrato 272
trompe de chasse 28
trompette marine 29
tuba 25, 45
typologie acoustique
formes sonores 157, 159, 167, 199
source vocale 443
sources instrumentales 297–311

V

verre
percuté, frotté 10, 62
rayonnement 48
vibraphone 11
vibrations 8
capteurs 18
inscription 33
matérialisation 19
vibrato musical 69, 266–271
amplitude (voix, flûte) 266
fréquence (violon, voix) 267
hauteur des notes courtes vibrées 271
largeur du vibrato (test) 268
Voir aussi modulations 121
vibrato vocal 42
émergence sur l'orchestre 225
vièle à roue (chien) 225
voix chantée
étude qualitative 380
formant du chanteur 373
voix chuchotée
anamorphose fréquentielle 168
émission 448
formes sémantiques 161, 451
intensité 219
résonances du conduit vocal 448
synthèse à l'icophone 168
voix de sifflet
voix humaine
cavités de résonance 451, 456
données anatomiques 445
écoutes 470
effet hélium 324
formants vocaliques 453
mécanismes laryngés 450
production du son 448–451
typologie acoustique 443–445
voyelles
chant harmonique 465
formants spectraux 455
hommes/femmes 453
intelligibilité 458
interaction son laryngé 457–459
qualité vocale 458, 459
quintina 469
tessiture vocale 457

Z

zone d'écoute dominante 116, 420
hauteur d'une cloche tube 261
schéma 276
voix de la quintina 469
zone sensible de l'oreille 208, 209

Index des noms propres

A

Andler, D. 150, 152
Aristote 140, 182, 395
Artaud, P.-Y. 337, 418
Assayag, G. 419
Asselin, P.-Y. 421, 423, 430, 432

B

Bachem, A. 437
Bacilly (de), B. 275
Bahnsen, P. 169
Bailly, L. 447, 464
Ballas, J. 148
Barbaras, R. 140
Barthélémy, J. P. 376
Bartók, B. 158, 292
Bartolozzi, B. 337, 418
Bayle, F. 369
Beauchamp, J. 287, 362
Bédos de Celles (Dom), F. 220, 289, 351, 427
Békésy (von), G. 81, 98
Bell, G. 140
Bennet, G. 174
Bensa, J. 364, 365
Bérard, J.-B. 289
Berg, A. 292
Berger, K. W. 311, 356
Berio, L. 294
Berlioz, H. 208, 290, 292, 337, 352, 353, 409
Berthoz, A. 143, 149, 180
Bertrand, D. 225
Besnainou, Ch. 3, 349, 378
Bettens, O. 424, 426
Bideaud, J. 182
Biot, J. M. 290
Bismarck (von), G. 356, 357, 358, 359
Blacking, J. 196, 201
Blauert, J. 131
Boehm, T. 276, 350
Bonnet, C. 190
Boring, E. 170, 180
Bosanquet, R. H. 426
Bouasse, H. 15, 418
Bougeret, G. 421, 432
Bouguer, P. 103
Boulez, P. 296, 301, 306
Brailoiu, C. 394
Bregman, A. 174, 175, 178, 201
Burns, E. 194, 392
Buser, P. 92, 94

C

Cage, J. 144, 293
Canévet, G. 135
Carpentier, G. 352
Caus (de), S. 395
Chaigne, A. 14, 15, 260, 304, 515
Chailley, J. 394, 395
Charpentier, M.-A. 430
Cheminée, P. 378
Cheveigné (de), A. 117
Chladni, E. 15, 19, 290
Chowning, J. 211
Clark, M. 214, 311, 326, 349, 356, 358
Clerc, P.-A. 430
Cogan, R. 171, 197, 299, 374
Cohen, E. 220
Cornut, G. 443, 449, 450
Couprie, P. 369
Cros, Ch. 74

D

D'Alessandro, C. 271
Dampierre (de), E. 395
Damsté, P. 446
Dargie, D. 464
Davis, H. 104
Debussy, C. 158, 293
Descartes, R. 22, 140
Deutsch, D. 174, 177, 178, 201
Donnadieu, S. 288, 356, 362, 377
Dubois, D. 182, 185, 188, 191, 192, 366, 367, 373, 376, 377, 378, 384, 385
Ducornet, M. 379
During, J. 400

E

Edison, T. 20, 33, 74
Ehrenfels (von), Ch. 141, 152
Eimert, H. 294
Engramelle, J. 220
Erickson, R. 171, 287, 293, 374, 384, 420
Escher, M. C. 170

F

Fabre, B. 337
Fabre, P. 449
Fant, G. 443
Faure, A. 346, 363, 377
Fechner, G. 47, 103, 140, 521
Féron, F.-X. 180
Fitch, W. T. 299
Fletcher, H. 15, 99, 108, 210
Fogliano, L. 396
Fokker, A. D. 426
Forest, L. 326, 350
Fourier, J. 22, 290, 353
Francès, R. 173, 230, 392
François, J.-Ch. 123, 260, 301
Fürniss, S. 394

G

Gaillard, P. 159, 363, 376
Ganassi, S. 220
Garcia (Jr), M. 446, 449
Garnier, M. 381
Gérard-Ardenois, D. 415
Gherghinoiou, C. 377
Gibson, J. 141
Globokar, W. 418
Godøy, R. I. 181
Grandville, J.-J. 180
Grey, J. M. 287, 358, 359, 360, 361, 362
Grisey, G. 73, 168, 293, 319, 415, 417
Guillaume, P. 140, 144, 148, 152, 153
Guyot, F. 188, 375, 377

H

Hajda, J. 288, 356, 362
Harvey, J. 73, 296, 346
Haydn, J. 408, 439
Haynes, B. 435
Helffer, M. 250, 299
Helmholtz (von), H. 19, 20, 33, 39, 74, 80, 98, 113, 124, 140, 180, 209, 288, 290, 305, 323, 325, 353, 354, 355, 398, 405, 415, 426, 529
Henrich, N. 381, 445, 451, 456, 464
Henry, P. 294, 369
Holbein, H. 167
Hotteterre, J. 220, 274, 275
Houdé, O. 182
Houtsma, A. J. M. 261
Huche, F. 443
Hugonnet, C. 135
Husserl, E. 118, 140, 146, 487

J

Jolas, B. 173
Jolivet, A. 337

K

Kaltenecker, M. 338
Kanizsa, G. 170
Kastner, G. 352
Kawahara, H. 168, 324
Kergomard, J. 14, 15, 260, 304, 350, 515
Kientzy, D. 418
Kircher, A. 525
Koechlin, Ch. 352, 353
Koenig, R. 74
Kondratov, A. 183
Kriehuber, J. 6
Krumhansl, C. 362

L

Labussière, A. 394
Lachenmann, H. 337
Lähdeoja, O. 18
Lakatos, S. 358, 362
Lakoff, G. 182
Lattard, J. 433
Lavoisier, A. 139
Le Nindre, B. 377
Lecocq, P. 147
Legaillard, Y. 430
Leipp, E. 29, 37, 51, 83, 125, 151, 162, 179, 208, 210, 224, 308, 335, 392, 409, 422, 435, 446, 515
Levinas, M. 337
Lévy, F. 420
Lichte, W. H. 356, 357, 358, 359
Liénard, J.-S. 456
Ligeti, G. 110, 222, 287, 293, 294
Lindley, M. 421, 431, 435
Lissajous, J.-A. 436
Lortat-Jacob, B. 470
Loulié, E. 351, 393
Luce, D. 326, 358

M

Mach, E. 299
Madelaine (de la), S. 273, 274, 487
Maffiolo, V. 191
Magritte, R. 180
Malherbe, C. 419
Manoury, P. 296
Marandas, E. 433
Martin, K. D. 311, 345
Mathews, M. 319, 358
Mattheson, J. 430
Mefano, P. 337
Merleau-Ponty, M. 129
Mersenne, M. 22, 28, 74, 115, 201, 351, 394, 396, 426, 431
Metfessel, M. 408
Miller, J. R. 358
Miller, R. 443
Miyazaki, K. 437
Moles, A. 146, 294
Molière 29, 452
Molino, J. 154
Moore, B. C. J. 99, 115, 117, 121, 125
Mori, T. 252, 434, 442
Mozart, W.-A. 271
Murail, T. 73
Musson-Gonnaud, V. 225

N

Neisser, U. 169, 185
Noorden (van), L. 175

O

Ormezzano, Y. 443
Ouzounoff, D. 418

P

Pascal, B. 7
Pateau, B. 418
Patterson, R. 94, 117, 299
Peeters, G. 345
Penrose, L. 242
Petitot, J. 152
Piaget, J. 182
Pierce, J. 15
Piéron, H. 103, 141
Plomp, R. 287, 356, 358
Praetorius, M. 435, 436, 443
Pressnitzer, D. 122, 123, 340, 416
Prokoviev, S. 339
Proust, M. 133
Psychoyou, T. 289, 351, 397, 421, 435

Q

Quantz, J.-J. 428

R

Rameau, J.-Ph. 221, 289, 290, 397
Rappoport, D. 415
Rayleigh, J. W. S. 130
Risset, J.-C. 180, 197, 205, 210, 242, 262, 287, 288, 295, 296, 356, 358, 420
Ritsma, R. J. 116, 261, 277
Rosch, E. 182, 183, 184
Rossing, T. 15, 264, 515
Roubeau, B. 450, 529
Rousseau, J.-J. 288, 289, 351, 416, 427
Roy, S. 295
Rubin, E. 170, 241
Russolo, L. 293

S

Sacks, O. 152
Saldanha, E. L. 311, 356
Sauveur, J. 15, 290, 396
Savart, F. 290
Scelsi, G. 293
Schaeffer, P. 154, 162, 190, 192, 287, 294, 295, 332, 369
Schafer, R. M. 369
Scherchen, H. 294
Schnupp, J. 127, 131
Schoenberg, A. 233, 292, 293
Schouten, J.F. 116, 257, 374
Scott de Martinville, E.-L. 20, 33, 74
Seashore, C. 408
Séchet, P. 326
Shamma, S. 90, 198
Shepard, R. 180, 242
Singer, R. 263
Slawson, W. 287, 374, 375

Sluchin, B. 418
Smalley, D. 369
Solomos, M. 296
Souris, A. 297, 300
Stevens, S. S. 104, 106, 109, 110, 128
Stockhausen, K. 293, 294
Strong, W. 356
Sundberg, J. 443

T
Tartini, G. 416
Tchaikovsky, P. 178
Tehrardt, E. 115
Thom, R. 157, 169
Titze, I. 443
Trân Quang Hai 246, 463
Trân Van Khe 29, 171
Traube, C. 375
Tromlitz, J. G. 428

V
Van Noorden, L. 174, 177
Vanderveer, N. 149, 154
Vangenot, S. 437
Varèse, E. 293, 294, 297, 337
Vennard, W. 443, 449, 450
Vicentino, N 425
Vogel, C. 3, 191

W
Warusfel, O. 133
Weber (von), C.-M. 417
Weber, E. H. 103
Weinreich, G. 133
Wertheimer, M. 141, 152, 174
Wessel, D. 174, 176, 177, 287, 291
Witkowski, N. 15
Wolfe, J. 324, 515
Wright, J. 27
Wu Run Shung 460

X
Xenakis, I. 106

Z
Zatorre, R. 437
Zwicker, E. 99, 122